现代物理基础丛书 46

实验数据分析

（下 册）

朱永生 著



《现代物理基础丛书》编委会

主　编　杨国桢

副主编　阎守胜　聂玉昕

编　委　（按姓氏笔画排序）

王　牧　王鼎盛　朱邦芬　刘寄星

邹振隆　宋菲君　张元仲　张守著

张海澜　张焕乔　张维岩　侯建国

侯晓远　夏建白　黄　涛　解思深



中国科学院科学出版基金资助出版

现代物理基础丛书　46

实验数据分析

(下册)

朱永生　著

科　学　出　版　社

北　京

内 容 简 介

本书介绍实验和测量数据分析中涉及的概率和数理统计及相关的数学知识，内容包括概率论、经典数理统计、贝叶斯统计、蒙特卡罗方法、极小化方法和去弥散方法六个部分. 特别讨论了数据统计处理中的一些困难问题和近期国际上发展起来的新方法. 书中分析了取自普通物理、核物理、粒子物理和工程技术问题的许多实例，注重物理问题与数学方法的结合，具体阐述了概率和数理统计及相关的数学方法在实际问题中的应用. 书末附有详尽的数理统计表，可供本书涉及的几乎所有数据分析问题之需要，而无需查阅专门的数理统计表书籍.

本书可供实验物理工作者和大专院校相关专业师生、理论物理研究人员、工程技术人员以及从事自然科学和社会科学的数据测量和分析研究人员参考.

图书在版编目（CIP）数据

实验数据分析. 下册/朱永生著. —北京：科学出版社, 2012
(现代物理基础丛书；46)
ISBN 978-7-03-034957-6
Ⅰ. ①实… Ⅱ. ①朱… Ⅲ. ①物理学-实验数据-分析 Ⅳ. ①O4-33
中国版本图书馆 CIP 数据核字(2012) 第 132863 号

责任编辑：钱　俊　鲁永芳／责任校对：包志虹
责任印制：钱玉芬／封面设计：陈　敬

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

中国科学院印刷厂印刷
科学出版社发行　各地新华书店经销
*
2012 年 6 月第　一　版　开本: B5(720 × 1000)
2012 年 6 月第一次印刷　印张: 24 1/2
字数: 480 000

定价：86.00 元
(如有印装质量问题, 我社负责调换)

前　言

自然科学和社会科学的许多领域, 诸如粒子物理和核物理、粒子和核天体物理、宇宙学和宇宙线物理、核工程和核医学、遗传学、人口统计、经济学、交通运输等等, 存在大量随机现象; 即使是对确定性现象的研究, 由于测量工具和仪器的测量误差通常具有随机性质, 所获得的数据也带有随机性. 因此, 实验或测量数据的分析几乎必定需要随机性数学及相关的数学分支的知识.

本书介绍实验或测量数据分析中所涉及的概率和数理统计及相关的数学知识, 内容包括概率论、经典数理统计、贝叶斯统计、蒙特卡罗方法、极小化方法和去弥散方法六个部分. 其中第 1∼5 章和第 6∼12 章分别阐述概率论和经典数理统计的基本内容, 第 13 章则专门介绍在现代统计学中具有重要影响的贝叶斯学派的观点与理论, 第 14 章讨论应用日益广泛的蒙特卡罗方法的基本概念, 第 15 章介绍的极小化 (或最优化) 方法是求解许多数理统计问题的重要工具 (例如, 极大似然法、最小二乘法等), 最后第 16 章介绍去弥散方法, 处理从观测数据和测量仪器的分辨函数反演出原分布的问题 (第 1∼11 章见本书上册).

一般的数学书籍侧重于数学上的严密和确切, 但对于以数学为工具的研究人员而言, 更为迫切的是了解相关数学的基本内容和方法, 并正确地运用它们. 因此, 本书力求数学与物理问题相结合, 避免过于抽象和过于数学化的讨论. 重点是介绍基本概念、基本原理和方法, 阐明方法的应用及适用条件, 而不是对定理作严格的证明和推导; 有些定理或结论只是直接引用, 但与数据处理直接相关的内容则予以充实. 本书中相当数量的例子取自普通物理、核物理、粒子物理和工程技术问题的数据分析问题的实际需要, 以期达到加深理解基本原理和正确运用它们的目的. 本书还力求反映国际上近期发展起来的处理数据分析中一些困难问题的新的概念和方法, 这方面的例子有第 11 章介绍的 “小信号测量的区间估计”、12.5 节讨论的 “信号的统计显著性”、8.8 节和 9.10 节讨论的 “多个实验结果的合并” 等内容. 书末附有详尽的数理统计表, 可供本书涉及的许多数据分析问题之需要, 而无需查阅专门的数理统计表书籍.

作者根据近年来在科研实践 (北京正负电子对撞机、北京谱仪的粒子物理实验研究) 和教学实践 (中国科学院研究生院讲授 “实验物理中的概率和统计” 课程) 中的体会, 认为有必要对科学出版社 2006 年出版的拙作《实验物理中的概率和统计 (第二版)》的内容作必要的修改和扩充. 增加的两章是近几十年来影响日益重要的贝叶斯统计, 以及在粒子物理和核物理、粒子和核天体物理、宇宙线物理和宇宙学、

光学和医学成像重建等领域得到广泛应用的去弥散方法. 在原有的章节中则增补了原书没有讨论但具有相当重要性的一些内容和概念, 如贝塔分布、对数正态分布和朗道分布, 费希尔信息, 区间估计的大样本法, 有约束的极大似然估计, 斯米尔诺夫-克拉美-冯　迈希斯检验, 多维随机变量分量的独立性检验, 相关性检验等. 此外, 还补充了与新增内容相对应的数理统计表, 增加了一些从实践中提炼出来的有典型意义的实例.

本书介绍了六部分共 16 章比较广泛的数学内容, 掌握这些知识能够大体应对数据分析对于随机性数学知识的基本需求, 至少可以提供进一步学习的基础. 事实上, 其中的每一部分都是一个重要的数学分支, 可以并且应当按照需要进行更为深入的学习. 对出于实用目的阅读本书的读者而言, 可以有选择地阅读部分章节. 书中第 1~5 章、第 6~10 章和第 12 章阐述概率论和经典数理统计的基本内容, 是利用随机性数学处理实际问题的基础; 第 11 章和第 13~16 章则可以根据需要加以选择.

本书的出版得到中国科学院科学出版基金的资助. 作者感谢中国工程院院士、中国高等科学技术中心学术主任叶铭汉研究员, 中国科学院研究生院郑阳恒教授, 中国科学院院士、中国科学院理论物理研究所张肇西研究员对于本书的热情鼓励和积极推介, 以及科学出版社钱俊等同志的大力支持和辛勤、细致的工作.

限于本人水平, 疏漏及不妥之处在所难免, 诚恳欢迎专家和读者批评指正.

朱永生

2011 年 9 月于北京

目　　录

前言
第 12 章　假设检验 401
12.1　假设检验的一般概念 401
12.1.1　原假设和备择假设 401
12.1.2　假设检验的一般方法 403
12.1.3　检验的比较 406
12.1.4　分布自由检验 408
12.2　参数假设检验 408
12.2.1　简单假设的奈曼–皮尔逊检验 408
12.2.2　复合假设的似然比检验 411
12.3　正态总体的参数检验 419
12.3.1　正态总体均值和方差的检验 419
12.3.2　两个正态总体均值的比较 421
12.3.3　两个正态总体方差的比较 423
12.3.4　多个正态总体均值的比较 427
12.4　拟合优度检验 429
12.4.1　似然比检验 430
12.4.2　皮尔逊 χ^2 检验 432
12.4.3　最小二乘、极大似然估计中的皮尔逊 χ^2 检验 435
12.4.4　拟合优度的一般 χ^2 检验 436
12.4.5　柯尔莫哥洛夫检验 443
12.4.6　斯米尔诺夫–克拉美–冯 · 迈希斯检验 447
12.5　信号的统计显著性 449
12.5.1　实验 P 值 449
12.5.2　信号的统计显著性 451
12.6　独立性检验 454
12.6.1　二维随机变量分量的独立性检验 454
12.6.2　多维随机变量分量的独立性检验 458
12.7　相关性检验 460

12.7.1 Pearson 相关系数的检验 ············ 460
12.7.2 Spearman 秩相关检验 ············ 463
12.7.3 Kendall τ 相关检验 ············ 466
12.7.4 多变量 Kendall 协和系数检验 ············ 471
12.8 一致性检验 ············ 474
12.8.1 符号检验 ············ 475
12.8.2 两子样的游程检验 ············ 480
12.8.3 游程检验作为皮尔逊 χ^2 检验的补充 ············ 484
12.8.4 两子样的斯米尔诺夫检验 ············ 487
12.8.5 两子样的威尔科克森检验 ············ 490
12.8.6 多个连续总体子样的克鲁斯卡尔–瓦列斯秩检验 ············ 495
12.8.7 多个离散总体子样的 χ^2 检验 ············ 498
第 13 章 贝叶斯统计 ············ 502
13.1 频率概率和贝叶斯概率 ············ 503
13.2 贝叶斯公式和贝叶斯统计模型 ············ 504
13.2.1 贝叶斯公式 ············ 504
13.2.2 贝叶斯统计模型和贝叶斯推断原则 ············ 506
13.2.3 先验分布和后验分布, 先验分布的选择 ············ 508
13.3 贝叶斯统计推断 ············ 527
13.3.1 统计决策的基本概念 ············ 527
13.3.2 贝叶斯参数点估计 ············ 531
13.3.3 经验贝叶斯估计 ············ 537
13.3.4 贝叶斯参数区间估计 ············ 540
13.3.5 贝叶斯假设检验 ············ 545
第 14 章 蒙特卡罗法 ············ 552
14.1 蒙特卡罗法的基本思想 ············ 552
14.2 随机数的产生及检验 ············ 554
14.2.1 随机数的产生 ············ 554
14.2.2 随机数的统计检验 ············ 556
14.3 任意随机变量的随机抽样 ············ 561
14.3.1 直接抽样方法 ············ 561
14.3.2 直接抽样方法的推广 —— 变换抽样 ············ 564
14.3.3 舍选抽样方法 ············ 567
14.3.4 利用极限定理抽样 ············ 569

14.3.5 复合分布的抽样方法 …… 570
14.3.6 近似抽样方法 …… 572
14.3.7 多维分布的抽样 …… 574
14.4 蒙特卡罗法计算积分 …… 580
14.4.1 频率法 (均匀投点法) …… 580
14.4.2 期望值估计法 …… 584
14.4.3 重要抽样方法 …… 587
14.4.4 半解析法 …… 588
14.4.5 自适应蒙特卡罗积分 …… 591
14.5 蒙特卡罗法应用于粒子传播问题 …… 593
第 15 章　极小化方法 …… 598
15.1 引言 …… 598
15.2 无约束极小化的一维搜索 …… 600
15.2.1 黄金分割法 (0.618 法) …… 601
15.2.2 斐波那契法 …… 603
15.2.3 二次函数插值法 (抛物线法) …… 607
15.2.4 进退法 …… 609
15.3 无约束 n 维极值的解析方法 …… 612
15.3.1 最速下降法 (梯度法) …… 613
15.3.2 牛顿法 …… 617
15.3.3 共轭方向法和共轭梯度法 …… 618
15.3.4 变尺度法 …… 624
15.4 无约束 n 维极值的直接方法 …… 626
15.4.1 坐标轮换法 …… 627
15.4.2 霍克–吉弗斯模式搜索法 …… 628
15.4.3 罗森布洛克转轴法 …… 629
15.4.4 单纯形法 …… 632
15.5 最小二乘 Q^2 函数和似然函数的极值问题 …… 635
15.5.1 最小二乘 Q^2 函数极值 …… 636
15.5.2 似然函数极值 …… 637
15.6 局部极小和全域极小 …… 639
15.6.1 网格法 …… 240
15.6.2 随机搜索法 …… 640
15.7 约束 n 维极值问题 …… 642
15.7.1 变量代换法 …… 643

15.7.2　罚函数法 ······ 644
15.8　参数的误差估计 ······ 648
第 16 章　去弥散方法 ······ 651
16.1　去弥散问题的数学表述 ······ 652
16.2　响应矩阵求逆法 ······ 656
16.3　修正因子法 ······ 660
16.4　正规化去弥散的一般策略 ······ 662
16.5　正规函数 ······ 663
16.5.1　Tikhonov 正规函数 ······ 663
16.5.2　基于极大熵原理的正规函数 ······ 665
16.5.3　贝叶斯统计的极大熵原理 ······ 666
16.5.4　基于交叉熵的正规函数 ······ 668
16.6　估计量的方差和偏差 ······ 669
16.7　正规参数的选择 ······ 672
16.8　去弥散计算实例 ······ 675
16.9　数值计算 ······ 678
参考文献 ······ 682
附表 ······ 691
示例索引 ······ 775
《现代物理基础丛书》已出版书目 ······ 779

第 12 章　假 设 检 验

12.1　假设检验的一般概念

从第 7 章到第 11 章我们讨论了参数估计问题. 在这类问题中, 随机变量的分布函数的形式一般为已知, 但其中包含着待估计的未知参数, 参数估计就是根据子样观测值对未知参数的数值或置信区间进行统计推断. 如果被观测的随机变量的分布函数的确切形式未知, 我们只能以假设的方式提出它所服从的分布, 并从统计的观点根据观测值来判断这一假设的合理性. 这类问题是数理统计的又一重要内容, 称为**统计假设的检验**.

举例来说, 方向相反的高能量正负电子对撞, 产生一对 μ 介子

$$\mathrm{e}^+ + \mathrm{e}^- \longrightarrow \mu^+ + \mu^-.$$

出射的 μ^- 粒子与负电子 e^- 之间的极角 ϑ 是一个随机变量. 假定测量了 n 个反应事例的 ϑ 值为 $\vartheta_1, \vartheta_2, \cdots, \vartheta_n$, 要求确定 ϑ 的分布是否具有

$$C(1 + a\cos^2\vartheta), \quad 0 \leqslant \vartheta \leqslant \pi \tag{12.1.1}$$

的形式, 其中 C 是归一化常数, a 是某个参数. 这就是一个假设检验问题.

假设检验可以分为**参数检验**和**非参数检验**两类. 如果有待检验的是分布的某个参数是否等于某个规定值 (分布函数形式已知, 但包含未知参数), 那么这属于参数检验问题. 比如上例中已知随机变量 ϑ 具有式 (12.1.1) 的分布, 要求根据观测值 $\vartheta_1, \vartheta_2, \cdots, \vartheta_n$ 检验未知参数 a 是否等于某个特定值 a_0. 非参数检验所处理的问题是：被观测的随机变量所服从的分布是否具有某个特定的函数形式, 或是从两个总体的各自一组观测值来检验这两个总体是否有相同的分布等, 在这种情况下, 待检验总体的分布的函数形式, 在假设检验完成前是无从知晓的. 上例中, 如果要根据一组观测值 $\vartheta_1, \vartheta_2, \cdots, \vartheta_n$ 来确定随机变量 ϑ 是否服从式 (12.1.1) 的分布 (事先并不知道 ϑ 分布的函数形式), 则就是非参数检验问题.

12.1.1　原假设和备择假设

参数检验的一般问题可表述如下：设总体 X 的概率分布 $F(x;\vartheta)$ 的函数形式为已知, 但其中包含未知参数 ϑ, 要求从总体的子样测量值 $(x_1, x_2, \cdots, x_n)$ 来检验

未知参数 ϑ 是否等于某个指定值 ϑ_0. 对我们要验证的假设记为

$$H_0 : \vartheta = \vartheta_0, \tag{12.1.2}$$

称为**原假设**或**零假设**. 参数假设检验问题的提出本身就意味着, 总体 X 的真实分布的参数值既可能是 H_0 规定的 ϑ_0, 也可能是不同于 ϑ_0 的其他值. 因此, 与原假设相对, 有

$$H_1 : \vartheta = \vartheta', \quad \vartheta' \neq \vartheta_0$$

称为**备择假设**或**备选假设**. 参数 ϑ 所有可能值的全体称为**容许假设**, 容许假设 (除原假设 H_0 以外) 都可作为备择假设. 常见的参数备择假设有如下类型:

$$H_1 : \vartheta = \vartheta_1 \quad (\vartheta_1\text{为不等于}\vartheta_0\text{的常数}), \tag{12.1.3}$$

$$H_1 : \vartheta > \vartheta_0, \tag{12.1.4}$$

$$H_1 : \vartheta < \vartheta_0, \tag{12.1.5}$$

$$H_1 : \vartheta \neq \vartheta_0. \tag{12.1.6}$$

如果假设对于参数的规定值是一个常数, 或者说是参数空间中的单点集, 则该假设称为**简单假设**; 相反, 假设对参数的规定值是参数空间中的非单点集, 则称为**复合假设**或**复杂假设**. 于是式 (12.1.2) 和式 (12.1.3) 是简单原假设和简单备择假设, 而式 (12.1.4)~ 式 (12.1.6) 是复合备择假设.

非参数检验的一类问题是, 待检验的总体 X 的分布 $F(x)$ 是否等于某个特定函数 $G(x)$, 或者总体 X 的分布 $F(x)$ 与总体 Y 的分布 $G(x)$ 是否相同. 其原假设可表述为

$$H_0 : F(x) = G(x), \tag{12.1.7}$$

备择假设可有不同的类型

$$H_1 : F(x) > G(x), \tag{12.1.8}$$

$$H_1 : F(x) < G(x), \tag{12.1.9}$$

$$H_1 : F(x) \neq G(x). \tag{12.1.10}$$

一个假设检验问题, 就是利用待检验总体的子样观测值来决定, 究竟应当接受原假设 (拒绝备择假设) 还是应当拒绝原假设 (接受备择假设). 至于原假设和备择假设怎样选择, 则是根据所要解决的具体问题来决定的.

式 (12.1.4), 式 (12.1.5) 的备择假设对于待检验的参数 ϑ 的规定值, 完全落在原假设 $\vartheta = \vartheta_0$ 的一侧 (上侧或下侧), 这样的检验称为**单侧检验**; 式 (12.1.6) 备择假设对 ϑ 的规定值落在 $H_0 : \vartheta = \vartheta_0$ 的两侧, 称为**双侧检验**. 对于非参数检验的情形, 式 (12.1.8), 式 (12.1.9) 是单侧检验, 式 (12.1.10) 是双侧检验.

12.1.2 假设检验的一般方法

设 $\boldsymbol{X}=\{X_1,X_2,\cdots,X_n\}$ 是从待检验总体抽取的随机子样, 而 $U=U(X)$ 为子样统计量 (见 6.2 节统计量的定义), 在假设检验中称为**检验统计量**, 令 W 是 U 的值域. 当零假设 H_0 为真时, U 落入 W 的一个子域 R 的概率用 α 表示, $0\leqslant\alpha\leqslant 1$,

$$\alpha=P(U\in R|H_0)=\int_R g(u|H_0)\mathrm{d}u, \tag{12.1.11}$$

其中 $g(u|H_0)$ 是 H_0 为真时统计量 U 的概率密度. 一般 α 为一接近于零的正数, 判断待检验的假设是拒绝还是接受, 是根据所谓小概率事件的原理, 即概率很小的事件在一次随机试验中被认为是几乎不可能发生的. 因此, 当我们有一组实际观测值 $x_1,\ x_2,\ \cdots,\ x_n$ 并求出 U 的实际观测值 U_{obs}, 如果它落在区域 R 之中, 由于 α 很小, 这一事件是小概率事件, 因此, 假设 H_0 不大可能是正确的, 我们称在**显著性(水平)**α 上拒绝零假设 H_0 而接受备选假设 H_1; 反之, 当 U_{obs} 落在子域 W–R 内, 则在水平 α 上接受 H_0 而拒绝 H_1, 对零假设 H_0 作出接受或拒绝的判断, 通常称为对 H_0 作**显著性检验**. 子域 R 称为**拒绝域**或**临界域**, 子域 W-R 则称为**接受域**. 临界域与接受域分界点的统计量 U 的值 U_{c} 称为**临界点**或**临界值**(图 12.1(a)). 应当指出, 在某些检验问题中, 特别在某些双侧检验问题中, 存在两个分隔开的临界域, 因而有两个临界点, 如图 12.1(b) 所示.

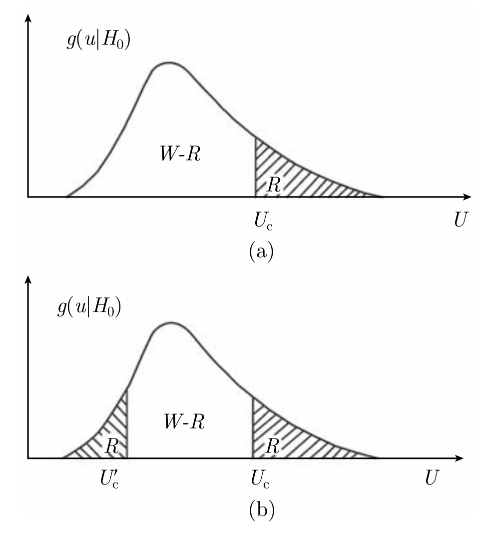


图 12.1 检验统计量 U 的临界域 R 和接受域 W-R

$U_{\text{c}}(U'_{\text{c}})$ 为临界值, $g(u|H_0)$ 是 H_0 为真时 U 的概率密度

由假设检验的上述判断准则可知, 即使零假设 H_0 为真, 但检验统计量 U 的实际观测值仍然有 α 的概率落入拒绝域 R, 也就是说, 当用 U_{obs} 来检验正确地反映

观测值的零假设时, 有 100α % 的可能性将拒绝 H_0. 这类错误称为**第一类错误**, 亦即**弃真**的错误, 把本来正确的假设给否定了. 为了减少弃真的错误, α 应当取得尽可能地小.

此外, 还可能出现**第二类错误**, 即**取伪的错误**, 当 H_0 不为真但却接受了 H_0. 出现取伪错误的概率取决于备择假设 H_1, 它等于 H_1 为真而 U 落入接收域 W-R 的概率 β,

$$\beta = P(U \in W - R|H_1) = \int_{W\text{-}R} g(u|H_1)\mathrm{d}u, \tag{12.1.12}$$

其中 $g(u|H_1)$ 表示 H_1 为真时统计量 U 的概率密度. 零假设 H_0 对备择假设 H_1 的**检验势**或**势函数**定义为

$$\text{检验势} = 1 - \beta = P(U \in R|H_1) = \int_R g(u|H_1)\mathrm{d}u, \tag{12.1.13}$$

即 H_1 为真而统计量 U 落入零假设拒绝域 R 的概率.

图 12.2 是假设检验中犯第一类错误的概率 α 和犯第二类错误的概率 β 的图示. 显然, 检验统计量 U 及临界值 U_c 的合理选择应当是使 α 尽可能地小, 使检验势 $1-\beta$ 尽可能大. 因而假设检验问题的症结在于选择适当的检验统计量 U 及其适当的临界值 U_c.

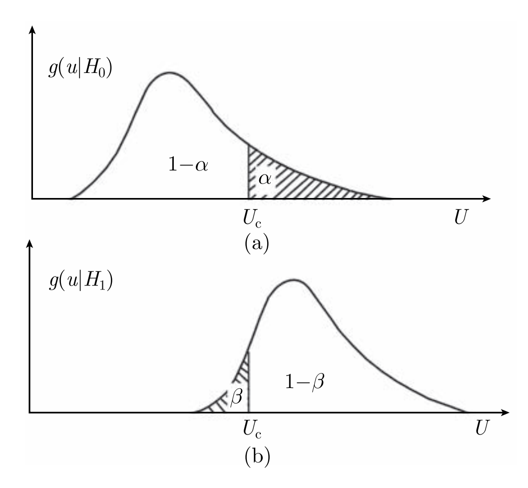


图 12.2 参数假设检验中第一类错误的概率 α 和第二类错误的概率 β

例 12.1 单个 π^0 和多个 π^0 事例的区分

考察在氢气泡室中质子反质子湮灭产生的粒子. 泡室只能显示带电粒子的径迹, 通过对径迹的测量可确定带电粒子的种类、飞行方向和动量; 中性粒子则不能显示和鉴别. $\mathrm{p\bar{p}}$ 反应的产物有许多事例观测到四条径迹, 并可鉴别出它们是 $\pi^\pm$ 介子. 但测定了这些 π 介子的动量后发现, 反应初态 $(\mathrm{p\bar{p}})$ 和反应末态 (4 个 π 介子)

之间不满足能量和动量守恒. 这表明, 反应末态中还有 "丢失" 了的中性粒子没有被观测到. 根据反应初态的能、动量和反应末态四个 π 介子的能、动量可以求出所谓的 "丢失质量"("丢失" 的中性粒子的静止能量之和), 事例数的丢失质量分布称为丢失质量谱. 分析丢失质量谱可知, 丢失的中性粒子可能是一个或多个中性 π^0 介子, 因此, $\mathrm{p\bar{p}}$ 反应事例可以分为产生一个 π^0 和产生多个 π^0 两类. 按照假设检验的概念, 现在的问题可用下述零假设和备择假设来表示：

$$
\begin{aligned}
&H_0 : \mathrm{p\bar{p}} \to \pi^+\pi^+\pi^-\pi^-\pi^0, \\
&H_1 : \mathrm{p\bar{p}} \to \pi^+\pi^+\pi^-\pi^- M \quad (M\text{表示多个}\pi^0),
\end{aligned}
$$

丢失质量的平方 m^2 作为检验统计量. 如果 H_0 成立, 即丢失了一个 π^0, 那么 m^2 应当等于 π^0 质量的平方, 即 $m^2_{\pi^0}$. 显然, 临界值 m^2_c 的合理选择应该是略高于 $m^2_{\pi^0}$. 这样, 如果一个事例的丢失质量平方小于 m^2_c, 就有很大可能是产生一个 π^0 的事例, 故接受 H_0 是合理的; 反过来若事例的丢失质量平方 m^2 大于 m^2_c, 那么有很大可能产生一个以上的 π^0, 故应当拒绝 H_0 而接受备择假设 H_1, 认为该事例是一个多 π^0 事件.

实验中观测到的全部事例的丢失质量谱一般都是连续分布, 例如, 图 12.3(a) 就是一个典型的丢失质量谱直方图. 这是一个实验分布, 其中包含了测量误差即实验分辨函数的效应 (见 4.17.1 节). 这样, 尽管真实的丢失质量小于 $m^2_{\pi^0}$, 但由于测量误差, 测得的 m^2 却有一定的概率大于 $m^2_{\pi^0}$; 反之, 真实的丢失质量大于 $m^2_{\pi^0}$ 时, 也有一定的概率实验测定值却小于 $m^2_{\pi^0}$. 这就模糊了单 π^0 事件与多 π^0 事件的界限, 使 m^2_c 的选择面临两难的境地. 如果 m^2_c 选得稍高于 $m^2_{\pi^0}$, 可以保证多 π^0 事例被误认为单 π^0 事例的概率很小, 即取伪错误的概率很小, 但真实的单 π^0 事例却有较大的可能损失掉 (弃真的概率较大); 反过来, 若 m^2_c 比 $m^2_{\pi^0}$ 大得多, 虽然减小了弃真错误的概率, 但取伪错误的概率却由此增大了. 这种情况在假设检验问题中是有代表性的, 减小 α 和减小 β 这两个要求常常互相抵触, 必须根据实际问题作适当的折中.

假定我们已知单 π^0 事例的 m^2 概率密度为 $f\left(m^2|H_0\right)$, 多 π^0 事例的 m^2 概率密度为 $f\left(m^2|H_1\right)$, 它们如图 12.3(b) 所示. 对于一个给定的显著性水平 α, 由式 (12.2.1) 得到 m^2 的临界值 m^2_c,

$$
\alpha = \int_{m^2_\mathrm{c}}^{\infty} f(m^2|H_0)\mathrm{d}m^2,
$$

代入式 (12.2.2), 求得 β,

$$
\beta = \int_0^{m^2_\mathrm{c}} f\left(m^2|H_1\right)\mathrm{d}m^2.
$$

计算表明, 当 α 由 0~0.1 时, $1-\beta$ 由 0 非常迅速地增大; 当 $\alpha \lesssim 0.15$, 检验势 $1-\beta$ 非常接近于 1. 如果希望单 π^0 事件的纯度高, 即多 π^0 事件的 "沾污" 很少, 则可取较大的 α; 如果要求少丢失单 π^0 事例, 则应取较小的 α(较高的 m_c^2). 可见, 临界值的选取要视问题的要求而定.

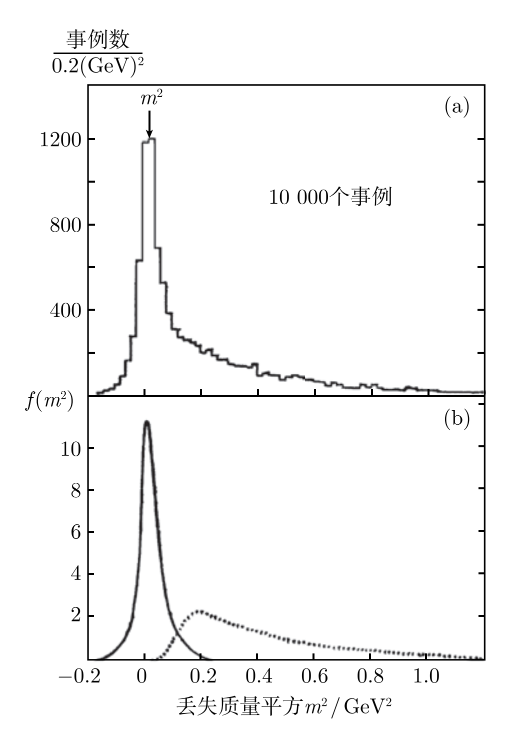


图 12.3　(a) $p\bar{p}$ 反应的丢失质量谱; (b) 丢失质量平方 m^2 的概率密度

实线 —— 单 π^0 事例; 虚线 —— 多 π^0 事例

12.1.3　检验的比较

12.1.2 节已经提到, 应当适当选择检验统计量及其临界值, 使得 α, β 同时尽可能地小. 因此, 不同检验统计量 (检验方法) 的优劣可以根据势函数 $p = 1 - \beta$ 来鉴定, 对于指定的 (相同) 显著性水平 α, 势函数越大, 检验方法越优越.

对于简单原假设、简单备择假设的参数检验问题:

$$H_0 : \vartheta = \vartheta_0, \quad H_1 : \vartheta = \vartheta_1.$$

由于 β 是统计量 U 概率密度 $g\,(u|H_1)$ 的函数, 故势函数可表示为

$$p(\vartheta_1) = 1 - \beta(\vartheta_1).$$

对于复合备择假设

$$H_1 : \vartheta \neq \vartheta_0,$$

有

$$p(\vartheta) = 1 - \beta(\vartheta).$$

于是在 $\vartheta = \vartheta_0$ 处, 有

$$p(\vartheta_0) = 1 - \beta(\vartheta_0) = \alpha.$$

图 12.4(a) 给出三种可能的势函数曲线, 它们对应于三种不同的检验方法 (检验统计量).

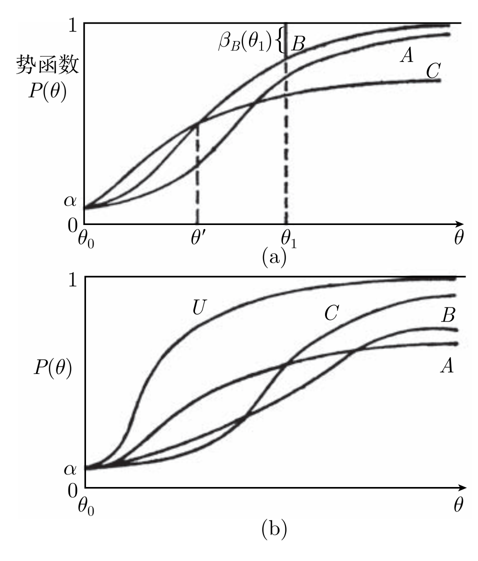


图 12.4 势函数和最佳势检验

对于简单原假设 $H_0 : \vartheta = \vartheta_0$ 和简单备择假设 $H_1 : \vartheta = \vartheta_1$ 的参数检验问题, 对于给定的显著性水平 α, 当势函数 $1 - \beta(\vartheta_1)$ 达到极大, 称为水平为 α 的**最佳势检验**, 简称为最佳势检验, 记为 MP(most powerful) 检验. 例如, 图 12.4(a) 中 B 检验是对 $H_1 : \vartheta = \vartheta_1$ 的 MP 检验. 如果在参数 ϑ 的某个子空间 D 中势函数 $1 - \beta(\vartheta)$, $\vartheta \in D$ 达到极大, 则称为**关于 D 的一致最佳势检验**. 记为关于 D 的 UMP(uniform most powerful) 检验. 这时, 备择假设可以是子空间 D 中的任意值或任意区间, 即可以是简单假设, 也可以是复合假设. 例如, B 检验是关于 $\vartheta > \vartheta'$ 的 UMP 检验, C 检验是 $(\vartheta_0, \vartheta')$ 区间内的 UMP 检验. 如果在 ϑ 的整个空间中势函数 $1 - \beta(\vartheta)$ 达到极大, 则是**一致最佳势检验**, 简称 UMP 检验. 图 12.4(b) 中, 任何检验统计量的势函数曲线对所有 ϑ 值都落在曲线 U 的下方, U 检验是 UMP 检验.

12.1.4　分布自由检验

在假设检验中, 核心问题是构造子样统计量 U. 知道了 U 的概率分布, 便可根据子样的测定值推断在显著性水平 α 上应接受还是拒绝原假设 H_0. 如果统计量 U 的概率分布并不显著地依赖于原假设 H_0 规定的待检验总体的分布函数的具体形式, 则称为检验是**分布自由**或**分布无关**的. 它的含义是, 只要原假设 H_0 为真, 不管 H_0 对待检验总体的分布规定的是什么函数形式, 统计量 U 的概率分布都相同, 并且是已知的. 因此, 对于给定的显著性水平 α, 临界域都相同. 这样, 对任意的原假设 H_0, 或者说对任意分布的总体, 都可使用这个检验统计量. 这就使得该检验方法有极大的普适性. 许多假设检验问题, 特别是非参数假设检验, 广泛地应用了分布自由的检验.

应当强调指出, 分布自由的性质只与检验的显著性水平相关联, 而不适用于检验的其他特征量. 特别是, 检验的势函数不具有分布自由的性质, 而是强烈地依赖于 H_0, H_1 对总体分布的具体规定.

本章的以后各节具体地讨论不同的假设检验问题. 其中 12.2 节, 12.3 节讨论参数假设检验, 12.4 节 ∼12.7 节则处理非参数假设检验.

12.2　参数假设检验

12.2.1　简单假设的奈曼 – 皮尔逊检验

首先我们来讨论简单原假设和简单备择假设这种最简单的情形. 令 ϑ 是待检验的总体分布中的未知参数, 于是有

$$H_0 : \vartheta = \vartheta_0, \qquad H_1 : \vartheta = \vartheta_1. \tag{12.2.1}$$

根据式 (12.1.13)H_0 对 H_1 的检验势可表示为

$$\begin{aligned} 1-\beta &= \int_R g(u|\vartheta_1)\mathrm{d}u = \int_R \frac{g(u|\vartheta_1)}{g(u|\vartheta_0)} \cdot g(u|\vartheta_0)\mathrm{d}u \\ &= \left[\frac{g(u|\vartheta_1)}{g(u|\vartheta_0)}\right]_{U=\xi} \cdot \int_R g(u|\vartheta_0)\mathrm{d}u = \alpha \left[\frac{g(u|\vartheta_1)}{g(u|\vartheta_0)}\right]_{U=\xi}, \end{aligned} \tag{12.2.2}$$

其中 ξ 是临界域中的某个点, $\left[\dfrac{g(u|\vartheta_1)}{g(u|\vartheta_0)}\right]_{U=\xi}$ 实际上是临界域中 $\dfrac{g(u|\vartheta_1)}{g(u|\vartheta_0)}$ 的某种平均值. 由以下事实:

$$\int_R \left[\frac{g(u|\vartheta_1)}{g(u|\vartheta_0)}\right]_{\min} g(u|\vartheta_0)\mathrm{d}u \leqslant \int_R \frac{g(u|\vartheta_1)}{g(u|\vartheta_0)} \cdot g(u|\vartheta_0)\mathrm{d}u \leqslant \int_R \left[\frac{g(u|\vartheta_1)}{g(u|\vartheta_0)}\right]_{\max} g(u|\vartheta_0)\mathrm{d}u,$$

立即可得到式 (12.2.2) 的结果. 由式 (12.2.2) 可知, 对于给定的显著性水平 α, 为了使势函数达到极大, 必须适当选择检验方法或检验统计量 U, 使得它的概率密度之比值 $\dfrac{g(u|\vartheta_1)}{g(u|\vartheta_0)}$ 在临界域 R 内尽可能地大. 这一原理称为**奈曼–皮尔逊**(Neyman-Pearson)**定理**.

使势函数达到极大的临界域称为**最佳临界域**, 它由满足下面不等式的点 (u 值) 组成：

$$\frac{g(u|\vartheta_1)}{g(u|\vartheta_0)} > k, \tag{12.2.3}$$

其中 k 是由给定的显著性 α 确定的常数.

如果检验统计量 U 取为观测值随机变量 X 自身, 那么 $g(u|\vartheta)$ 就是观测值所服从的总体的概率密度 $f(x|\vartheta)$. 当进行一系列测量得到观测值 $X = \{X_1, X_2, \cdots, X_n\}$, 概率密度函数 $f(x|\vartheta)$ 需用联合概率密度即似然函数代替

$$L(X|\vartheta) = \prod_{i=1}^{n} f(X_i|\vartheta),$$

判据式 (12.2.3) 于是需改写为 (最佳临界域由满足下面不等式的 x 值组成) :

$$\frac{L(X|\vartheta_1)}{L(X|\vartheta_0)} > k, \tag{12.2.4}$$

这时, 显著性水平 α 为

$$\int_R L(\boldsymbol{X}|\vartheta_0)\mathrm{d}\boldsymbol{X} = \alpha. \tag{12.2.5}$$

按照式 (12.2.3)∼ 式 (12.2.5) 建立的临界域, 对于给定的显著性水平 α, 将给出简单零假设对于简单备择假设的最佳势检验 (势函数值达到极大), 该临界域称为**最佳临界域**. 它使得犯第二类错误 (取伪的错误) 的概率达到极小.

对于多次测量的情形, 观测值随机变量的最佳临界域很难找到, 因为式 (12.2.5) 是一个 n 维积分, 有时很难求出. 在实际问题中, 应当寻找适当的检验统计量 (观测量的某个函数) 的最佳临界域. 例如, 当对总体的均值 α 进行检验时, 比较方便的是寻求子样平均 $\overline{X}$ 的临界域; 而检验总体方差 σ^2 则可寻求子样方差 s^2 的临界域等. 这时应对所使用的统计量的概率密度函数作积分.

例 12.2 粒子平均寿命的奈曼–皮尔逊检验

假定不稳定粒子的衰变时间 t 有 n 个观测值 $t_1,\ t_2, \cdots,\ t_n$. 粒子衰变时间 t 的概率密度形式为

$$f(t|\tau) = \frac{1}{\tau}\exp\left(\frac{-t}{\tau}\right), \tag{12.2.6}$$

其中 τ 为粒子平均寿命. 要求对下列简单零假设和备择假设进行检验:

$$H_0 : \tau = 1, \quad H_1 : \tau = 2. \tag{12.2.7}$$

根据式 (12.2.4), 最佳临界域应满足

$$\frac{L(\boldsymbol{t}|\tau=2)}{L(\boldsymbol{t}|\tau=1)} = \frac{\prod\limits_{i=1}^{n} \frac{1}{2}\exp\left(-\frac{t_i}{2}\right)}{\prod\limits_{i=1}^{n}\exp(-t_i)} = \left(\frac{1}{2}\right)^n \exp\left(\frac{1}{2}\sum_{i=1}^{n} t_i\right) > k.$$

利用子样平均的记号 $\bar{t} = \frac{1}{n}\sum\limits_{i=1}^{n} t_i$, 上式可写成

$$\bar{t} > 2\left(\frac{1}{n}\ln k + \ln 2\right) \equiv T_n, \tag{12.2.8}$$

即在 $\bar{t}$ 的空间中, 最佳临界域是满足不等式 (12.2.8) 的一组值, 其中 T_n 是一个由选定的显著性 α 决定的常数. 我们知道, 平均寿命 τ 是粒子衰变时间 t 的期望值, 用子样平均 $\bar{t}$ 作为检验统计量是十分自然的. 为了找出最佳临界域对应的 T_n 值, 必须知道 $\bar{t}$ 的概率密度 $f_n(\bar{t})$. 对于 $n=1$ 和 n 很大这两种极限情形, $f_n(\bar{t})$ 的形式特别简单. 下面, 我们选定显著性水平 H_0= 5%, 就这两种情形分别加以讨论.

(1) n=1. 这时 $\bar{t}$ 与 t 的概率密度相同

$$f_1(\bar{t}) = \frac{1}{\tau}\exp\left(\frac{-\bar{t}}{\tau}\right).$$

$\bar{t}$ 的临界域的下限 T_1 可由显著性 α 的定义 (式 (12.1.11)) 确定

$$0.05 = \alpha = \int_{T_1}^{\infty} \mathrm{e}^{-\bar{t}}\mathrm{d}\bar{t}, \quad T_1 = -\ln\alpha \approx 3.00.$$

于是零假设 H_0:n=1 对备择假设 H_1:$\tau = 2$ 的检验势由式 (12.1.13) 得到

$$1-\beta = \int_{T_1}^{\infty} \frac{1}{2}\mathrm{e}^{-\bar{t}/2}\mathrm{d}\bar{t} = \sqrt{\alpha} \approx 0.22.$$

在实验中, 对粒子衰变时间作一次测量, 如果测量值大于 $T_1 = 3.00$, 即落入临界域, 则拒绝零假设 H_0. 但是在这种检验中, 取伪的错误 (H_1 为真却接受 H_0) 的可能性很大, 其概率由 β 表示, $\beta \approx 0.78$.

(2) n 很大. 这种情形下检验统计量 $\bar{t}$ 的概率密度可用均值 τ, 方差 τ^2/n 的正态函数作为近似 (见 8.3.6 节)

$$f_n(\bar{t}) = N(\tau, \ \tau^2/n) = \frac{1}{\sqrt{2\pi}\tau/\sqrt{n}}\exp\left[-\frac{1}{2}\frac{(\bar{t}-\tau)^2}{\tau^2/n}\right].$$

$\bar{t}$ 的临界域的下限 T_n 由下式求得：

$$
\begin{aligned}
0.05 = \alpha &= \int_{T_n}^{\infty} N\left(1,\ \frac{1}{n}\right)\mathrm{d}\bar{t} = 1 - \int_{-\infty}^{T_n} N\left(1,\ \frac{1}{n}\right)\mathrm{d}\bar{t} \\
&= 1 - \varPhi\left(\frac{T_n - 1}{1/\sqrt{n}}\right),
\end{aligned}
$$

其中 $\varPhi$ 是累积标准正态函数. 查表可得

$$
T_n = 1 + \frac{1.645}{\sqrt{n}}.
$$

$H_0 : \tau = 1$ 对于 $H_1 : \tau = 2$ 的检验势则为

$$
\begin{aligned}
1 - \beta &= \int_{T_n}^{\infty} N\left(2,\ \frac{4}{n}\right)\mathrm{d}\bar{t} = 1 - \varPhi\left(\frac{T_n - 2}{2/\sqrt{n}}\right) \\
&= 1 - \varPhi\left(0.8225 - \frac{\sqrt{n}}{2}\right).
\end{aligned}
$$

因此, 临界值 T_n 和检验势 $1-\beta$ 两者都有赖于观测次数 n. 若取 $n = 100$, 则有 $T_{100} = 1.16, 1 - \beta = 0.999\,99$. 由势函数 $1-\beta$ 的表达式可见, 当观测次数 n 趋于无穷, 势函数趋向 1, 即取伪的错误概率等于 0, 这称为相合检验 (consistent test).

由以上讨论可以看出, 对于给定的显著性水平 $\alpha = 0.05$, 随着观测数 n 的增加 $(n = 1 \to 100)$, H_0 对 H_1 的检验势迅速增加 $(1 - \beta = 0.22 \to 0.999\,99)$; 如果取较小的 α 值 (减小弃真错误的概率), 例如, $\alpha = 0.01$, 当 $n = 1 \to 100$, 检验势的值下降为 $1 - \beta = 0.10 \to 0.999\,4$.

12.2.2 复合假设的似然比检验

奈曼–皮尔逊检验仅适用于 H_0 和 H_1 都是简单假设的情形, 如果 H_0 和 H_1 中至少有一个是复合假设, 则必须用似然比检验.

设随机变量 X 的概率密度为 $f(x|\boldsymbol{\vartheta})$, 未知参数 $\boldsymbol{\vartheta} = \{\vartheta_1, \cdots, \vartheta_k\}$, $\boldsymbol{\vartheta} \in \varOmega, \varOmega$ 为参数空间. 假定零假设 H_0 是对 $\vartheta_1, \cdots, \vartheta_k$ 中至少一个参数加上某个约束条件 (如等于某个常数), 使得 $\boldsymbol{\vartheta}$ 被限制在参数空间 $\varOmega$ 的一个子空间 ω 中. 我们的问题是根据总体 X 的容量 n 的子样 $\boldsymbol{X} = \{X_1, \cdots, X_n\}$ 来检验假设

$$
H_0 : \boldsymbol{\vartheta} \in \omega, \quad H_1 : \boldsymbol{\vartheta} \in \varOmega - \omega. \tag{12.2.9}
$$

对于子样值 $\boldsymbol{X} = \{X_1, \cdots, X_n\}$, 似然函数为

$$
L = \prod_{i=1}^{n} f(x_i|\boldsymbol{\vartheta}).
$$

记似然函数在参数空间 Ω 中的极大值为 $L(\hat{\Omega})$, 而在零假设 H_0 为真的条件下得到的子空间 ω 中似然函数的极大值记为 $L(\hat{\omega})$. **似然比**λ 定义为

$$\lambda \equiv \frac{L(\hat{\omega})}{L(\hat{\Omega})} . \tag{12.2.10}$$

由概率密度的非负性可知, λ 为非负值; 同时, 在子空间 ω 中的极大值 $L(\hat{\omega})$ 不可能大于整个参数空间 Ω 中的极大值 $L(\hat{\Omega})$, 故有 $0 \leqslant \lambda \leqslant 1$.

似然比 λ 是观测值 $x_1, x_2, \cdots, x_n$ 的函数. 若 λ 的观测值接近于 1, 那么 H_0 为真时的极大值 $L(\hat{\omega})$ 与整个参数空间的极大值 $L(\hat{\Omega})$ 相接近, 这表示 H_0 为真的可能性很大; 反之, 若 λ 很小, 则 H_0 为真的可能性很小. 直观地就可知道, 似然比 λ 是零假设 H_0 的合理的检验统计量.

似然比检验的方法可陈述如下**:** 令 H_0 为真时, 似然比 λ 的概率密度为 $g(\lambda|H_0)$, 对于给定的显著性水平 α, λ 的临界域由

$$0 < \lambda < \lambda_\alpha \tag{12.2.11}$$

确定, 其中 λ_α 满足

$$\alpha = \int_0^{\lambda_\alpha} g(\lambda|H_0) \mathrm{d}\lambda. \tag{12.2.12}$$

如果 λ 的观测值 λ_{obs} 大于 λ_α, 则在水平 α 上接受零假设 H_0; 反之, 则拒绝 H_0.

如果函数 $g(\lambda|H_0)$ 为未知, 只要知道 λ 的某个单调函数的概率密度, 仍然可以进行似然比检验. 设 $y = y(\lambda)$ 是 λ 的单调函数, y 的概率密度 $h(y|H_0)$ 为已知, 则有

$$\alpha = \int_0^{\lambda_\alpha} g(\lambda|H_0) \mathrm{d}\lambda = \int_{y(0)}^{y(\lambda_\alpha)} h(y|H_0) \mathrm{d}y, \tag{12.2.13}$$

求出变量 y 的临界域

$$y(0) < y < y(\lambda_\alpha), \tag{12.2.14}$$

通过 λ 与 $y(\lambda)$ 的逆变换容易求出临界值 λ_α.

通常似然比 λ(或它的函数) 的严格分布是很难找到的, 这时, H_0 的检验问题相当复杂, 一般可以尝试采用近似方法.

不失一般性, 假设式 (12.2.9) 通常可写成如下形式

$$\begin{aligned}
H_0: \quad & \vartheta_0 = \vartheta_{i0}, && i = 1, 2, \cdots, r, && (\text{也可表示为}\boldsymbol{\vartheta}_r = \boldsymbol{\vartheta}_{r0}); \\
& \vartheta_j\text{不限定}, && j = 1, 2, \cdots, s, && (\text{也可表示为}\boldsymbol{\vartheta}_s\text{不限定}). \\
H_1: \quad & \vartheta_i \neq \vartheta_{i0}, && i = 1, 2, \cdots, r, && (\text{也可表示为}\boldsymbol{\vartheta}_r \neq \boldsymbol{\vartheta}_{r0}); \\
& \vartheta_j\text{不限定}, && j = 1, 2, \cdots, s, && (\text{也可表示为}\boldsymbol{\vartheta}_s\text{不限定}).
\end{aligned} \tag{12.2.15}$$

这里原假设限定了参数集的一个子集的数值 ($\boldsymbol{\vartheta}_r = \boldsymbol{\vartheta}_{r0}$), 而备择假设 H_1 中这一子集可取 H_0 限定的参数值之外的任意值. 于是, 现在式 (12.2.10) 定义的似然比是固定 $\boldsymbol{\vartheta}_r = \boldsymbol{\vartheta}_{r0}$ 而对 $\boldsymbol{\vartheta}_s$ 求极大的 $L(\boldsymbol{X}|\boldsymbol{\vartheta})$ 值与对于全部参数求极大的 $L(\boldsymbol{X}|\boldsymbol{\vartheta})$ 值之比. 利用上述记号, 似然比可写为

$$\lambda = \frac{L\left(\boldsymbol{X}|\boldsymbol{\vartheta}_{r0}, \boldsymbol{\vartheta}_s''\right)}{L\left(\boldsymbol{X}|\boldsymbol{\vartheta}_r', \boldsymbol{\vartheta}_s'\right)}, \tag{12.2.16}$$

式中 $\boldsymbol{\vartheta}_s''$ 是原假设 H_0 限定的参数空间中的极大似然函数对应的 $\boldsymbol{\vartheta}_s$ 值, 而 $\boldsymbol{\vartheta}_r', \boldsymbol{\vartheta}_s'$ 是全部参数空间 $\boldsymbol{\vartheta}$ 中的极大似然函数对应的 $\boldsymbol{\vartheta}_r, \boldsymbol{\vartheta}_s$ 值.

可以证明, 当 H_0 为真时, 如子样容量 n 很大, 则统计量 $-2\ln\lambda$ 趋近于 $x^2(r)$ 分布[35]. 或者更一般地, 如果 H_0 对于全部参数空间 $\boldsymbol{\vartheta}$ 中的的 $s+r$ 个参数施加了 r 个约束, 则 $-2\ln\lambda$ 当 H_0 为真时渐近地服从 $\chi^2(r)$ 分布. 因此, 可利用统计量 $-2\ln\lambda$ 来检验零假设 H_0, 通过 $\chi^2(r)$ 分布概率密度函数的积分来确定 λ 的临界域.

例 12.3 正态分布均值的似然比检验

设总体分布为正态函数

$$N(\theta_1, \theta_2) = \frac{1}{\sqrt{2\pi\vartheta_2}} \exp\left[-\frac{1}{2}\frac{(x-\vartheta_1)^2}{\vartheta_2}\right], \quad \vartheta_2 > 0,$$

给定 n 个观测值 $X_1, X_2, \cdots, X_n$, 要求检验其分布的均值是否等于常数 μ_0.

检验的零假设和备择假设可表为

$$H_0: \vartheta_1 = \mu_0,$$
$$H_1: \vartheta_1 \neq \mu_0.$$

参数空间 $\varOmega$ 是 $\vartheta_2 > 0$ 的半平面中所有点的集合, 即

$$-\infty < \vartheta_1 < \infty, \quad \vartheta_2 > 0.$$

当零假设 H_0 成立, 参数 $\vartheta_1 = \mu_0$, 故子空间 ω 是 $\vartheta_1 = \mu_0, \vartheta_2 > 0$ 的一条半无界直线. n 个观测值的似然函数为

$$L(\boldsymbol{X}|\vartheta_1, \vartheta_2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\vartheta_2}} \exp\left[-\frac{1}{2}\frac{(X_i-\vartheta_1)^2}{\vartheta_2}\right]. \tag{12.2.17}$$

参数空间 $\varOmega$ 中 ϑ_1, ϑ_2 的极大似然估计量可由 8.2 节 (3) 的方法求出

$$\hat{\vartheta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}, \tag{12.2.18}$$

$$\hat{\vartheta}_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2. \tag{12.2.19}$$

代入式 (12.2.17), 得到似然函数在 Ω 中的极大值

$$L(\hat{\Omega}) = \left(\frac{n}{2\pi\sum\limits_{i=1}^{n}(X_i - \overline{X})^2}\right)^{n/2} \mathrm{e}^{-n/2}. \tag{12.2.20}$$

在子空间 ω 中, 似然函数的极大值可将 $\vartheta_1 = \mu_0$ 代入式 (12.2.17) 再对 ϑ_2 求极大得到, 其结果是

$$\hat{\vartheta}_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2. \tag{12.2.21}$$

这一表达式与式 (12.2.19) 不同, 即在参数空间 Ω 和 ω 中, ϑ_2 的极大似然估计量是不同的. 这样, 当 H_0 为真, 似然函数极大值由式 (12.2.21) 代入式 (12.2.17) 得出

$$L(\hat{\omega}) = \left(\frac{n}{2\pi\sum\limits_{i=1}^{n}(X_i - \mu_0)^2}\right)^{n/2} \mathrm{e}^{-n/2}. \tag{12.2.22}$$

根据似然比 λ 的定义, 我们有

$$\lambda = \left(\frac{\sum\limits_{i=1}^{n}(X_i - \overline{X})^2}{\sum\limits_{i=1}^{n}(X_i - \mu_0)^2}\right)^{n/2}. \tag{12.2.23}$$

因为 $\sum\limits_{i=1}^{n}(X_i - \mu_0)^2 = \sum\limits_{i=1}^{n}(X_i - \overline{X})^2 + n(\overline{X} - \mu_0)^2$, 所以 λ 可以写为

$$\lambda = \left(\frac{1}{1 + \dfrac{t^2}{n-1}}\right)^{n/2}, \tag{12.2.24}$$

$$t = \frac{\sqrt{n}(\overline{X} - \mu_0)}{\sqrt{\sum\limits_{i=1}^{n}(X_i - \overline{X})^2 \Big/ (n-1)}} = \frac{\dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)s^2}{\sigma^2} \Big/ (n-1)}}. \tag{12.2.25}$$

由 6.3.3 节所述可知, 式 (12.2.25) 中的 t 正是自由度 $(n-1)$ 的 t 分布变量, 其概率密度 $f(t;\ n-1)$ 在 4.18 节中给定. 由式 (12.2.24) 可求得 t 与 λ 的函数关系 $t=t(\lambda)$, 于是变量 λ 的概率密度 $g(\lambda)$ 可从 t 分布导出, 从而按式 (12.2.13) 求出给定显著性 α 时 λ 的临界域. 但这一变量代换并不必须实行, 因为从式 (12.2.24) 可知, λ 是 t^2 的单调函数. 当 H_0 不为真时 λ 很小, t^2 很大, 因此, $0<\lambda<\lambda_\alpha$ 所确定的临界域对应于 $t^2>t_\alpha^2$. 这样, 对于给定的显著性水平 α, t 的临界域由两个子区间构成

$$-\infty<t<-t_{\alpha/2}, \quad t_{\alpha/2}<t<\infty. \tag{12.2.26}$$

$t_{\alpha/2}$ 定义为

$$\frac{\alpha}{2}=\int_{-\infty}^{-t_{\alpha/2}} f(t;\ n-1)\mathrm{d}t \equiv \int_{t_{\alpha/2}}^{\infty} f(t;\ n-1)\mathrm{d}t. \tag{12.2.27}$$

如果观测值 $x_1, x_2, \cdots, x_n$ 对应的 t 值 t_{obs} 满足

$$-t_{\alpha/2}<t_{\mathrm{obs}}<t_{\alpha/2},$$

则接受零假设 H_0; 否则 t_{obs} 落入式 (12.2.26) 定义的临界域, H_0 被拒绝.

例如, 考察一个子样容量 $n=20$ 的实验, 选定显著性 $\alpha=0.05$. 由附表 8 查得 $t_{\alpha/2}=t_{0.025}=2.093$. 如果由式 (12.2.25) 算出的 $|t_{\mathrm{obs}}|$ 值大于 2.093, 则应当拒绝零假设 H_0. 将 $t_{0.025}=2.093$ 代入式 (12.2.24), 可求出对应的 λ 临界值为 $\lambda_{0.05}=0.125$.

我们将 λ 临界值的上述似然比计算与本节中提到的近似计算作一比较. 按近似方法, 统计量 $-2\ln\lambda$ 近似地服从 χ^2 分布, 自由度等于 H_0 中取确定值的参数个数, 在现在的情形中, 自由度等于 1. 根据式 (12.2.23) 可知

$$-2\ln\lambda=n\ln\left[1+\frac{n(\overline{X}-\mu_0)^2}{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}\right].$$

将上式右边的对数作幂级数展开并只取第一项, 用 σ^2 代替子样方差, 得出近似表达式

$$-2\ln\lambda\approx\left(\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right)^2.$$

对于 $\chi^2(1)$ 变量, 显著性 $\alpha=0.05$ 对应的临界值为 $-2\ln\lambda_{0.05}=3.841$, $\lambda_{0.05}=0.147$. 与似然比方法求出的数值 $\lambda_{0.05}=0.125$ 对比可见, 对于子样容量小到 $n=20$ 的场合, $-2\ln\lambda$ 近似还是合理的.

为了计算检验势 $1-\beta$, 必须考虑备择假设 $H_1: \vartheta_1 \neq \mu_0$. 式 (12.2.25) 定义的变量仍作为检验统计量, 但当 H_1 为真时, t 表达式中的

$$\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$$

不再是 $N(0,1)$ 分布, 而是 $N(\mu,1)$ 正态变量

$$\sqrt{\delta} \equiv \mu \equiv \frac{\vartheta_1-\mu_0}{\sigma/\sqrt{n}}. \tag{12.2.28}$$

所以统计量 t 服从自由度 $n-1$ 的非中心 t 分布, 非中心参数为 δ (见 4.18 节). 于是势函数为

$$\begin{aligned}1-\beta(\delta) &= \int_0^{\lambda_\alpha} g(\lambda|H_1,\delta)\mathrm{d}\lambda \\ &= \int_{-\infty}^{-t_{\alpha/2}} f(t;\ n-1,\delta)\mathrm{d}t + \int_{t_{\alpha/2}}^{\infty} f(t;\ n-1,\delta)\mathrm{d}t,\end{aligned}$$

其中 $g(\lambda|H_1,\delta)$ 是 H_1 为真时对于给定的 δ 值似然比 λ 的概率密度, $f(t;\ n-1,\ \delta)$ 是自由度 $n-1$、非中心参数 δ 的非中心 t 分布概率密度. 参考文献 [95] 可查到 $f(t;\ n-1,\ \delta)$ 的累积分布函数值. 图 12.5 画出了不同的 α, n 值对应的势函数. 由图可知, 本问题中的势函数有以下特点：

(1) 势函数对 $\delta=0$ 为对称, 在 $\delta=0$ 处达到极小, 随着 $|\delta|$ 的增加而增加.

(2) 势函数随着显著性 α 的增加而增加.

(3) 对于给定的 δ 值, 势 $1-\beta$ 随着子样容量 n 的增加而增加.

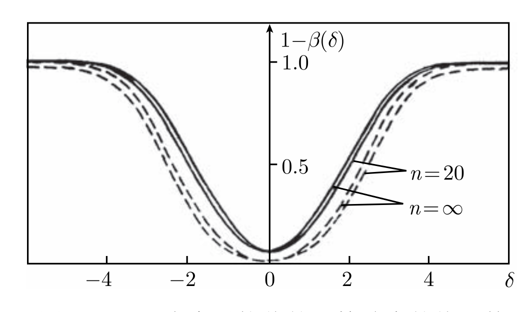


图 12.5　正态变量均值的 t 检验中的势函数

显著性 $\alpha=0.05$(实线) 和 0.02(虚线), 子样容量 n=20 和 $n=\infty$

值得指出, 本问题中 λ 的临界域由

$$0<\lambda<\lambda_\alpha$$

确定, 因为 $0 \leqslant \lambda \leqslant 1$, 所以临界域 $[0, \lambda_\alpha]$ 处于 λ 值域的下侧, 称为**下侧临界域**, 相应地 λ_α 称为**下侧临界值**. 而对于变量 t, 它的临界域由式 (12.2.26) 确定, 分别处于 t 值域的上、下两侧, 存在**双侧临界域**, 相应地存在**上侧临界值**$t_{\alpha/2}$ 和**下侧临界值**$-t_{\alpha/2}$. 双侧和单侧临界域在以后的讨论中经常会遇到.

例 12.4 粒子平均寿命的似然比检验

现在考虑例 12.2 讨论过的粒子平均寿命问题. 原假设仍为简单假设, 但备择假设是复合假设

$$H_0 : \tau = 1, \quad H_1 : \tau \neq 1.$$

对于子样值 $\boldsymbol{t} = \{t_1, t_2, \cdots, t_n\}$, 似然函数是

$$L = \prod_{i=1}^{n} \frac{1}{\tau} \exp\left(\frac{-t_i}{\tau}\right).$$

由 8.1 节知, τ 的极大似然估计为 $\bar{t}$. 似然函数中 τ 代之以 $\bar{t}$ 即得到 τ 的值域中似然函数的极大值

$$L(\hat{\Omega}) = \prod_{i=1}^{n} \frac{1}{\bar{t}} \exp\left(\frac{-t_i}{\bar{t}}\right) = \frac{1}{(\bar{t})^n} \exp\left\{\sum_{i=1}^{n} \frac{-t_i}{\bar{t}}\right\} = (\bar{t})^{-n} \exp(-n).$$

零假设为真时, $\tau = 1$, 代入似然函数, 即得

$$L(\hat{\omega}) = \prod_{i=1}^{n} \exp(-t_i) = \exp\left(\sum_{i=1}^{n} -t_i\right) = \exp(-n\bar{t}),$$

于是似然比为

$$\lambda = \frac{L(\hat{\omega})}{L(\hat{\Omega})} = (\bar{t})^n \exp\left[-n(\bar{t} - 1)\right].$$

当 n 很大, 统计量 $\bar{t}$ 的概率密度近似于正态分布 (见例 12.2):

$$f_n(\bar{t}) \sim N\left(\tau,\ \frac{\tau^2}{n}\right).$$

根据式 (12.2.13), 当取显著性水平 $\alpha = 0.05$, 应有

$$\begin{aligned} 0.05 = \alpha &= \int_{\bar{t}_0(\lambda=0)}^{\bar{t}_\alpha(\lambda=\lambda_\alpha)} f_n\left(\bar{t} | H_0\right) \mathrm{d}\bar{t} = \int_{\bar{t}_0}^{\bar{t}_\alpha} N\left(1,\ \frac{1}{n}\right) \mathrm{d}\bar{t} \\ &= \int_{-\infty}^{\bar{t}_\alpha} N\left(1, \frac{1}{n}\right) \mathrm{d}\bar{t} = \varPhi\left[\sqrt{n}(\bar{t}_\alpha - 1)\right]. \end{aligned}$$

查表得到 $\sqrt{n}\,(\bar{t}_\alpha - 1) = 1.645$, 于是得出临界值

$$\begin{aligned} \bar{t}_\alpha &= 1 + \frac{1.645}{\sqrt{n}}, \\ \lambda_\alpha &= \left(1 + \frac{1.645}{\sqrt{n}}\right)^n \exp\left(-1.645\sqrt{n}\right). \end{aligned}$$

当 λ 的观测值 λ_{obs} 小于 λ_α, 或 $\bar{t}$ 的观测值 $\bar{t}_{\text{obs}}$ 小于 $\bar{t}_\alpha$, 则在显著性水平 α=0.05 上拒绝零假设. 显然, 当 $n\to\infty$, $\bar{t}_\alpha$ 和 $\lambda_\alpha\to 1$.

例 12.5 粒子衰变理论的似然比检验

设试图通过某一实验来确定某个粒子两个衰变道复衰变振幅的比值 $\boldsymbol{X}$：

$$\boldsymbol{X}=\frac{\text{衰变1振幅}}{\text{衰变2振幅}}=\vartheta_{\text{Re}}+\mathrm{i}\vartheta_{\text{Im}}.$$

一般情形下 $\boldsymbol{X}$ 可以是任意复数, 但是存在三种理论对 X 有如下预期值：

假设 A：当理论 A 正确, 则 $\vartheta_{\text{Re}}=\vartheta_{\text{Im}}=0$.

假设 B：当理论 B 正确, 则 ϑ_{Re} 为非 0 实数, $\vartheta_{\text{Im}}=0$.

假设 C：当理论 C 正确, 则 ϑ_{Im} 为非 0 实数, $\vartheta_{\text{Re}}=0$.

就当前而言, 我们感兴趣的只是在假设 A,B,C 或一般情形之间作出鉴别. 因此这里我们所关心的主要是不同假设之间的鉴别, 而不甚关心 $\boldsymbol{X}$ 的具体值.

该实验中, 假设 A 是简单假设, 假设 B 和 C 是复合假设. 与所有这些假设相对立的备择假设是 $\text{Re}(\boldsymbol{X})$ 和 $\text{Im}(\boldsymbol{X})$ 均不为 0, 即一般情形亦为复合假设, 所以必须利用似然比检验.

根据一组实验数据计算了对数似然函数 $\ln L(\boldsymbol{X})$, 并且在 $\text{Re}(\boldsymbol{X})$ 和 $\text{Im}(\boldsymbol{X})$ 空间中画出似然函数极大值邻域的等值线, 如图 12.6 所示. 其中 $\boldsymbol{X}=d$ 是似然函数极大值点, 而 $\boldsymbol{X}=b$ 和 $\boldsymbol{X}=c$ 分别对应于 $\text{Im}(\boldsymbol{X})$=0 和 $\text{Re}(\boldsymbol{X})$=0 时的似然函数极大值点.

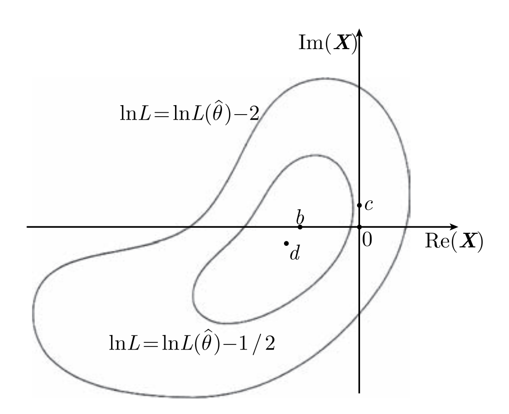


图 12.6 衰变振幅实验中对数似然函数值的等位线及约束极大值

原假设 A 相对于一般情形 (备择假设) 的似然比为

$$\lambda_a=\frac{L(0)}{L(d)}.$$

如果假设 A 为真, 则 $-2\ln\lambda_a$ 渐近地服从 $\chi^2(2)$ 分布, 这里自由度为 2, 因为原假设对于两个参数 $\vartheta_{\text{Re}},\vartheta_{\text{Im}}$ 都施加了约束.

原假设 B 相对于一般情形的极大似然比为

$$\lambda_b = \frac{L(b)}{L(d)}.$$

如果假设 B 为真, 则 $-2\ln\lambda_b$ 渐近地服从 $\chi^2(1)$ 分布, 自由度为 1 是因为原假设仅对于一个参数 ϑ_{Im} 施加了约束.

类似地, 原假设 C 相对于一般情形的极大似然比为

$$\lambda_c = \frac{L(c)}{L(d)}.$$

如果假设 C 为真, 则 $-2\ln\lambda_c$ 渐近地服从 $\chi^2(1)$ 分布.

因此, 根据 $-2\ln\lambda_a, -2\ln\lambda_b, -2\ln\lambda_c$ 的实验观测值可给出对于理论 A、B、C 的检验.

12.3　正态总体的参数检验

本节集中讨论正态总体的参数检验问题, 包括单个正态总体均值、方差的检验, 两个以上正态总体参数比较的检验等. 重点在于研究怎样构造适当的检验统计量.

12.3.1　正态总体均值和方差的检验

设 $X_1, X_2, \cdots, X_n$ 是正态总体 $N(\mu, \sigma^2)$ 的一个随机子样, 我们要从这组子样值来确定分布的均值是否为某一确定值 μ_0, 故零假设和备择假设分别是

$$H_0: \mu = \mu_0, \quad H_1: \mu \neq \mu_0. \tag{12.3.1}$$

这类检验问题的处理与 7.7 节中求正态总体均值的置信区间的过程十分相像. 检验零假设 $H_0: \mu = \mu_0$ 的适当检验统计量取决于总体方差 σ^2 是否为一已知量. 由 7.7 节所述可知

$$\sigma^2\text{已知}, \quad Z \equiv \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1),$$

$$\sigma^2\text{未知}, \quad Z \equiv \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1). \tag{12.3.2}$$

故 Z 可作为检验统计量. 式中 $\overline{X}$ 和 S^2 是子样平均和子样方差

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

式 (12.3.1) 显然是一个双侧检验问题. 对于给定的显著性水平 α, 可利用标准正态分布 (σ^2 已知) 或 t 分布 (σ^2 未知) 的概率密度的积分来确定检验统计量的双侧临界值, 即使

$$\frac{\alpha}{2}=\int_{-\infty}^{z_d} f(z)\mathrm{d}z=\int_{z_u}^{\infty} f(z)\mathrm{d}z \tag{12.3.3}$$

成立, 则 z_d 和 z_u 分别为下侧临界值和上侧临界值, 其中 $f(z)$ 是变量 Z 的概率密度, 对于 σ^2 为已知和未知时分别为 $N(0,1)$ 和 $t(n-1)$ 的概率密度, 由于它们对 0 点为对称, 故有

$$z_d=-z_u.$$

如果 Z 的观测值 Z_{obs} 满足

$$z_d<Z_{\mathrm{obs}}<z_u,$$

即 Z_{obs} 落入接受域, 则在显著性 α 上接受零假设; 否则就拒绝 H_0.

正态总体方差的检验可采用类似于 7.8 节所描述的方法. 零假设和备择假设分别是

$$H_0:\sigma^2=\sigma_0^2,\quad H_1:\sigma^2\neq\sigma_0^2. \tag{12.3.4}$$

由 7.8 节可知

$$\mu\text{已知},\quad Z\equiv\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}\sim\chi^2(n),$$

$$\mu\text{未知},\quad Z\equiv\frac{(n-1)s^2}{\sigma_0^2}=\frac{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}\sim\chi^2(n-1). \tag{12.3.5}$$

因而 Z 可作为检验统计量. 利用式 (12.3.3)($f(z)$ 现在是 $\chi^2(n)$ 和 $\chi^2(n-1)$ 的概率密度) 可求出双侧临界值 z_d 和 z_u. 由于 χ^2 是非对称分布, 故 $z_d\neq-z_u$.

如果备选假设的形式为

$$H_1:\sigma^2>\sigma_0^2\quad\text{或}\quad\sigma^2<\sigma_0^2,$$

则称为单侧检验, 其 z_d 或 z_u 由下式确定：

$$\alpha=\int_{-\infty}^{z_d} f(z)\mathrm{d}z\quad(\text{下侧临界值}), \tag{12.3.6}$$

$$\alpha=\int_{z_u}^{\infty} f(z)\mathrm{d}z\quad(\text{上侧临界值}). \tag{12.3.7}$$

12.3.2　两个正态总体均值的比较

设 $X_1, X_2, \cdots, X_n$ 是正态总体 $N(\mu_1, \sigma_1^2)$ 容量 n 的子样, $Y_1, Y_2, \cdots, Y_m$ 是 $N(\mu_2, \sigma_2^2)$ 容量 m 的子样. 希望检验两个正态总体是否有相同的均值, 因此有

$$H_0: \mu_1 = \mu_2, \quad H_1: \mu_1 \neq \mu_2. \tag{12.3.8}$$

这类检验对应的实际问题是, 希望知道两组实验测量数据实际上是否同一个物理量的测量值. 这一问题看起来似乎很简单, 实际上, 除非对方差 σ_1^2 和 σ_2^2 作某些假定, 否则很难作检验. 我们来考虑如下三种情形:

(1) σ_1^2 和 σ_2^2 为已知

由正态分布的性质知道, 子样平均 $\overline{X}$ 和 $\overline{Y}$ 服从正态分布 $N(\mu_1, \sigma_1^2/n)$ 和 $N(\mu_2, \sigma_2^2/m)$. 按照正态变量的加法原理, $\overline{X} - \overline{Y}$ 也是正态变量, 其均值和方差分别为 $\mu_1 - \mu_2, \dfrac{\sigma_1^2}{n} + \dfrac{\sigma_2^2}{m}$. 于是

$$\frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}} \sim N(0, 1). \tag{12.3.9}$$

为了检验 $\mu_1 - \mu_2 = 0$ 的零假设 H_0, 适当的检验统计量为

$$Z \equiv \frac{\overline{X} - \overline{Y}}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}}. \tag{12.3.10}$$

若 H_0 为真, 它应是 $N(0, 1)$ 变量. 这是方差已知的正态分布均值检验问题, 12.3.1 节已经加以讨论.

例如, 两组实验的结果分别表示为 $\bar{x} \pm \Delta\bar{x}$ 和 $\bar{y} \pm \Delta\bar{y}$, n_x 和 n_y 是两组实验的子样容量 (测量次数), 假定两组测量都服从正态分布, 那么 $\Delta\bar{x}$ 和 $\Delta\bar{y}$ 分别等于 $\sqrt{\sigma_x^2/n_x}$ 和 $\sqrt{\sigma_y^2/n_y}$. 当检验两个实验是否测量了同一物理量时, 适当的检验统计量为

$$Z \equiv \frac{\bar{x} - \bar{y}}{\sqrt{(\Delta\bar{x})^2 + (\Delta\bar{y})^2}}. \tag{12.3.11}$$

如果 H_0 为真 (测量的是同一物理量), 则 Z 应服从 $N(0, 1)$. 式 (12.3.11) 的 Z 是物理学家利用实验结果来表示的检验统计量的方便形式.

(2) σ_1^2 和 σ_2^2 未知但相等

由 6.3.1 节和 6.3.2 节知道, 利用 $N(\mu_1, \sigma_1^2)$ 的子样 $X_1, X_2, \cdots, X_n$ 和 $N(\mu_2, \sigma_2^2)$ 的子样 $Y_1, Y_2, \cdots, Y_m$, 可得到四个相互独立且分布已知的变量

$$\overline{X} \sim N(\mu_1, \sigma_1^2/n), \quad \overline{Y} \sim N(\mu_2, \sigma_2^2/m),$$

$$\frac{(n-1)S_1^2}{\sigma_1^2} \sim \chi^2(n-1), \quad \frac{(m-1)S_2^2}{\sigma_2^2} \sim \chi^2(m-1).$$

我们利用这四个变量来构造检验统计量. 两个正态变量 $\overline{X}, \overline{Y}$ 可构成标准正态变量

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}} \sim N(0,1);$$

根据 χ^2 分布的加法原理, 两个 χ^2 变量之和产生一个新的 χ^2 变量

$$\frac{(n-1)S_1^2}{\sigma_1^2}+\frac{(m-1)S_2^2}{\sigma_2^2} \sim \chi^2(n+m-2).$$

上述两个 $N(0,1)$ 变量和 $\chi^2(n+m-2)$ 变量也相互独立, 它们可按式 (4.18.4) 的方式构成一个 t 变量

$$\frac{\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}}}{\sqrt{\left[(n-1)s_1^2/\sigma_1^2+(m-1)s_2^2/\sigma_2^2\right]/(n+m-2)}} \sim t(n+m-2). \tag{12.3.12}$$

其中包含未知参数 σ_1^2 和 σ_2^2 因而不能作为检验统计量. 但当 $\sigma_1^2=\sigma_2^2$ 时, 它可进一步简化为

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}} \sim t(n+m-2), \tag{12.3.13}$$

其中 S 定义为

$$\begin{aligned} S^2 &\equiv \frac{1}{n+m-2}\left[(n-1)S_1^2+(m-1)S_2^2\right] \\ &= \frac{1}{n+m-2}\left[\sum_{i=1}^{n}(X_i-\overline{X})^2+\sum_{i=1}^{m}(Y_i-\overline{Y})^2\right]. \end{aligned} \tag{12.3.14}$$

它是总体方差 $\sigma^2(=\sigma_1^2=\sigma_2^2)$ 的估计.

因此, 当检验零假设 $H_0: \mu_1=\mu_2$ 时, 取

$$Z \equiv \frac{\overline{X}-\overline{Y}}{S\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}} \tag{12.3.15}$$

作为检验统计量. 若 H_0 为真, 则 Z 服从自由度 $n+m-2$ 的 t 分布. 于是问题化为 12.3.1 节中讨论过的方差未知的正态总体均值的检验问题.

从以上推导过程中可以看到, 除非两个正态总体的未知方差相等, 一般无法求出具有已知分布的变量作为检验统计量. 因此, $\sigma_1^2 \neq \sigma_2^2$ 的一般情形无法严格求解, 只能求助于近似方法.

(3) σ_1^2, σ_2^2 未知且不相等

在这种情形下, 若子样容量 n, m 足够大, 子样方差可以作为总体方差的估计. 在式 (12.3.9) 中用子样方差 S_1^2, S_2^2 代替 σ_1^2, σ_2^2, 则有

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{S_1^2/n+S_2^2/m}} \sim N(0,1). \tag{12.3.16}$$

当检验零假设 $H_0: \mu_1 = \mu_2$ 时, 取

$$Z \equiv \frac{\overline{X}-\overline{Y}}{\sqrt{S_1^2/n+S_2^2/m}} \tag{12.3.17}$$

作为检验统计量. 若 H_0 为真, 则近似地 Z 服从标准正态分布 $N(0,1)$. 问题化为方差已知的正态分布均值的检验问题.

12.3.3　两个正态总体方差的比较

设两个正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 的各自的子样是 $X_1, X_2, \cdots, X_n$ 和 $Y_1, Y_2, \cdots, Y_m$, 要求检验它们是否有相同的方差. 于是零假设和备择假设为

$$H_0: \sigma_1^2 = \sigma_2^2, \quad H_1: \sigma_1^2 \neq \sigma_2^2. \tag{12.3.18}$$

首先讨论两个总体均值 μ_1, μ_2 未知的情形. 已经知道

$$\frac{(n-1)S_1^2}{\sigma_1^2} = \frac{1}{\sigma_1^2}\sum_{i=1}^{n}(X_i-\overline{X})^2 \sim \chi^2(n-1),$$

$$\frac{(m-1)S_2^2}{\sigma_2^2} = \frac{1}{\sigma_2^2}\sum_{i=1}^{m}(Y_i-\overline{Y})^2 \sim \chi^2(m-1),$$

这两个变量相互独立, 由 4.19 节可知

$$\frac{\dfrac{(n-1)S_1^2}{\sigma_1^2}\Big/(n-1)}{\dfrac{(m-1)S_2^2}{\sigma_2^2}\Big/(m-1)} = \frac{S_1^2}{S_2^2}\cdot\frac{\sigma_2^2}{\sigma_1^2} \sim F(n-1, m-1). \tag{12.3.19}$$

因此, 当 H_0 成立, 则有

$$Z \equiv \frac{S_1^2}{S_2^2} = \frac{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2\Big/(n-1)}{\sum\limits_{i=1}^{m}(Y_i-\overline{Y})^2\Big/(m-1)} \sim F(n-1, m-1), \tag{12.3.20}$$

变量 Z 不含任何未知参数, 而且分布已知, 可作为检验统计量 ·

如果两个总体的期望值 μ_1,μ_2 都已知, 当 H_0 为真, 即 $\sigma_1^2=\sigma_2^2=\sigma^2$, 则 $\sum_{i=1}^{n}\frac{(X_i-\mu_1)^2}{\sigma^2}\sim\chi^2(n)$, $\sum_{i=1}^{m}\frac{(Y_i-\mu_2)^2}{\sigma^2}\sim\chi^2(m)$, 即有

$$Z\equiv\frac{\sum_{i=1}^{n}(X_i-\mu_1)^2\Big/n}{\sum_{i=1}^{m}(Y_i-\mu_2)^2\Big/m}\sim F(n,m). \tag{12.3.21}$$

注意, 此时统计量 Z 服从自由度 n,m 的 F 分布.

总结 12.3.1 节 ~12.3.3 节正态总体的参数检验问题, 将不同检验的有关结论汇总于表 12.1.

例 12.6　两个测量装置测量结果的比较

设动量 p_0=24.90GeV/c 的单能粒子束流射入泡室, 利用两个测量装置 A 和 B 测量 20 条粒子径迹在磁场中的曲率半径, 从而确定粒子的动量 p. 测到的 $1/p$ 的实验数据见表 12.2. 假定 $1/p$ 是一近似的正态变量, 径迹动量的测定误差仅仅由装置的误差决定, 但数值未知.

首先我们想知道, 两个装置的测量值 $\langle 1/p\rangle$ 是否与粒子动量 p_0 的倒数相一致. 这是一个方差未知的正态分布的均值检验问题, 即

$$H_0:\frac{1}{p}=\frac{1}{p_0},\quad H_1:\frac{1}{p}\neq\frac{1}{p_0}.$$

按表 12.1, 这属于编号 2-1 的检验问题. 合适的检验统计量为

$$t=\frac{\left\langle\frac{1}{p}\right\rangle-\frac{1}{p_0}}{S\big/\sqrt{20}},$$

当 H_0 为真, $t\sim t(19)$. 选定显著性水平 $\alpha=0.05$, 从附表 8 查得双侧检验的临界域是

$$|t|>t_{0.025}=2.09.$$

将上表中的数据代入 t 的表达式, 得到两个实验的观测值分别为

$$t_A=-0.38,\quad t_B=1.57.$$

这两个值都位于接受域内. 因此可以认为, 在 $\alpha=0.05$ 的显著性水平上, 两台装置对粒子动量的测定是合理的.

表 12.1 正态总体均值和方差的检验

检验参数	检验假设 H_0	检验假设 H_1	条件	检验统计量及分布	临界域	编号
μ	$\mu = \mu_0$	$\mu \neq \mu_0$	σ^2 已知	$U \equiv \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$	$\|U\| > z_{\alpha/2}$	1-1
	$\mu \geqslant \mu_0$	$\mu < \mu_0$			$U < -z_\alpha$	1-2
	$\mu \leqslant \mu_0$	$\mu > \mu_0$			$U > z_\alpha$	1-3
	$\mu = \mu_0$	$\mu \neq \mu_0$	σ^2 未知	$T \equiv \frac{\overline{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1)$	$\|T\| > t_{\alpha/2}(n-1)$	2-1
	$\mu \geqslant \mu_0$	$\mu < \mu_0$			$T < -t_\alpha(n-1)$	2-2
	$\mu \leqslant \mu_0$	$\mu > \mu_0$			$T > t_\alpha(n-1)$	2-3
	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$\sigma_1^2 = \sigma_2^2 = \sigma^2$ σ^2 已知	$U \equiv \frac{\overline{X}-\overline{Y}}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim N(0,1)$	$\|U\| > z_{\alpha/2}$	3-1
	$\mu_1 \geqslant \mu_2$	$\mu_1 < \mu_2$			$U < -z_\alpha$	3-2
	$\mu_1 \leqslant \mu_2$	$\mu_1 > \mu_2$			$U > z_\alpha$	3-3
	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$\sigma_1^2 \neq \sigma_2^2$ σ_1^2, σ_2^2已知	$U \equiv \frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}} \sim N(0,1)$	$\|U\| > z_{\alpha/2}$	4-1
	$\mu_1 \geqslant \mu_2$	$\mu_1 < \mu_2$			$U < -z_\alpha$	4-2
	$\mu_1 \leqslant \mu_2$	$\mu_1 > \mu_2$			$U > z_\alpha$	4-3
	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$\sigma_1^2 = \sigma_2^2 = \sigma^2$ σ^2 未知	$T \equiv \frac{\overline{X}-\overline{Y}}{S\sqrt{1/n+1/m}} \sim t(n+m-2)$ $S^2 \equiv \frac{1}{n+m-2}\left[(n-1)S_1^2+(m-1)S_2^2\right]$	$\|T\| > t_{\alpha/2}(n+m-2)$	5-1
	$\mu_1 \geqslant \mu_2$	$\mu_1 < \mu_2$			$T < -t_\alpha(n+m-2)$	5-2
	$\mu_1 \leqslant \mu_2$	$\mu_1 > \mu_2$			$T > t_\alpha(n+m-2)$	5-3
	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$\sigma_1^2 \neq \sigma_2^2$ σ_1^2, σ_2^2未知	$U \equiv \frac{\overline{X}-\overline{Y}}{\sqrt{\frac{S_1^2}{n}+\frac{S_2^2}{m}}} \sim N(0,1)$	$\|U\| > z_{\alpha/2}$	6-1
	$\mu_1 \geqslant \mu_2$	$\mu_1 < \mu_2$			$U < -z_\alpha$	6-2
	$\mu_1 \leqslant \mu_2$	$\mu_1 > \mu_2$			$U > z_\alpha$	6-3

续表

检验参数	检验假设		条件	检验统计量及分布	临界域	编号
	H_0	H_1				
σ^2	$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	μ已知	$\chi^2 \equiv \sum\limits_{i=1}^{n} \frac{(X_i-\mu)^2}{\sigma_0^2} \sim \chi^2(n)$	$\chi^2 \notin \left[\chi^2_{1-\alpha/2}(n), \chi^2_{\alpha/2}(n)\right]$	7-1
	$\sigma^2 \geqslant \sigma_0^2$	$\sigma^2 < \sigma_0^2$			$\chi^2 < \chi^2_{1-\alpha}(n)$	7-2
	$\sigma^2 \leqslant \sigma_0^2$	$\sigma^2 > \sigma_0^2$			$\chi^2 > \chi^2_{\alpha}(n)$	7-3
	$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	μ未知	$\chi^2 \equiv \frac{(n-1)S^2}{\sigma_0^2} = \sum\limits_{i=1}^{n} \frac{(X_i-\overline{X})^2}{\sigma_0^2} \sim \chi^2(n-1)$	$\chi^2 \notin \left[\chi^2_{1-\alpha/2}(n-1), \chi^2_{\alpha/2}(n-1)\right]$	8-1
	$\sigma^2 \geqslant \sigma_0^2$	$\sigma^2 < \sigma_0^2$			$\chi^2 < \chi^2_{1-\alpha}(n-1)$	8-2
	$\sigma^2 \leqslant \sigma_0^2$	$\sigma^2 > \sigma_0^2$			$\chi^2 > \chi^2_{\alpha}(n-1)$	8-3
	$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	μ_1, μ_2已知	$F \equiv \frac{\sum\limits_{i=1}^{n}(X_i-\mu_1)^2/n}{\sum\limits_{i=1}^{m}(Y_i-\mu_2)^2/m} \sim F(n,m)$	$F \notin \left[F_{1-\frac{\alpha}{2}}(n,m), F_{\frac{\alpha}{2}}(n,m)\right]$	9-1
	$\sigma_1^2 \geqslant \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$			$F < F_{1-\alpha}(n,m)$	9-2
	$\sigma_1^2 \leqslant \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$			$F > F_{\alpha}(n,m)$	9-3
	$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	μ_1, μ_2未知	$F \equiv \frac{S_1^2}{S_2^2} = \frac{\sum\limits_{i=1}^{n}(X_i-\bar{X})^2/n}{\sum\limits_{i=1}^{m}(Y_i-\bar{Y})^2/m} \sim F(n-1,m-1)$	$F \notin \left[F_{1-\frac{\alpha}{2}}(n-1,m-1), F_{\frac{\alpha}{2}}(n-1,m-1)\right]$	10-1
	$\sigma_1^2 \geqslant \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$			$F < F_{1-\alpha}(n-1,m-1)$	10-2
	$\sigma_1^2 \leqslant \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$			$F > F_{\alpha}(n-1,m-1)$	10-3

表中 $z_{\alpha/2}$ 为标准正态分布双侧 α 分位数

表 12.2 装置 A, B 测定的粒子动量数据

装置	$\left\langle \frac{1}{p} \right\rangle = \frac{1}{20}\sum_{i=1}^{20}\frac{1}{p_i}$ $10^{-3}(\text{GeV}/c)^{-1}$	$Y \equiv \left[\frac{1}{20}\sum_{i=1}^{20}\left(\frac{1}{p_i}-\frac{1}{p_0}\right)^2\right]^{1/2}$ $10^{-3}(\text{GeV}/c)^{-1}$
A	40.12	0.46
B	40.25	0.25

其次我们想了解两台装置的测量精度是否一致. 于是待检验的假设为

$$H_0: \frac{1}{\sigma_A^2} = \frac{1}{\sigma_B^2}, \quad H_1: \frac{1}{\sigma_A^2} \neq \frac{1}{\sigma_B^2}.$$

这里 σ 表示装置对粒子动量测量的精度. 已经知道这两个正态分布的均值是 $\mu_A = \mu_B = 1/p_0$, 根据表 12.1, 该检验问题属于编号 9-1 的检验问题, 检验统计量是

$$F = Y_A^2/Y_B^2.$$

当 H_0 为真, 它服从自由度 (20, 20) 的 F 分布. 选定显著性水平 $\alpha = 0.05$, 由附表 9 查得 $f_{0.975} = 2.46$, 再由 4.16 节的式 (4.16.11) 求得 $f_{0.025} = 0.406\,5$, 因此, 假设 H_0 的接受域为 [0.4065, 2.46]. 由观测值 Y_A 和 Y_B 知

$$f_{\text{obs}} = \frac{0.46^2}{0.25^2} = 3.39,$$

f_{obs} 落在 H_0 的临界域内, 故在显著性水平 $\alpha = 0.05$ 上拒绝两台装置精度相等的零假设, 并由 $Y_A > Y_B$ 知道装置 B 对粒子动量的测量精度比较高.

12.3.4 多个正态总体均值的比较

设有 N 个实验对同一个物理量进行观测, 得到的测量值及其误差表示为 $X_i \pm \Delta X_i, i = 1, \cdots, N$. 其中每个实验的测量值 X_i 通常是一组观测值的平均值, ΔX_i 是平均值的误差. 待检验的问题是, 这 N 个测量结果是否在误差范围内一致. 假定每个实验中观测服从正态分布, 于是这是一个 N 个正态总体均值的检验问题, 零假设可表示为

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_N, \tag{12.3.22}$$

备择假设 H_1 为任何其他的可能性.

如果 H_0 规定了 $\mu_1 = \cdots = \mu_N = \mu$ 的数值, 并且每个实验中的方差 σ_i^2 为已知, 那么变量

$$\sum_{i=1}^{N} \frac{(X_i - \mu)^2}{\sigma_i^2} \tag{12.3.23}$$

不含任何未知参数, 而且服从 $\chi^2(N)$ 分布 (见式 (4.14.21)) 故可以作为检验统计量. 但现在 σ_i 和 μ 都为未知量, 必须用实验观测值来作为近似. ΔX_i 可作为 σ_i 的估计, 而总体均值 μ 的估计值可取为 N 次测量的加权平均

$$\hat{\mu} = \overline{X} = \frac{\sum_{i=1}^{N} w_i X_i}{\sum_{i=1}^{N} w_i}, \tag{12.3.24}$$

μ 的误差的估计值则可取为

$$\hat{\sigma} = \Delta\overline{X} = \left(\sum_{i=1}^{N} w_i\right)^{-1/2}, \tag{12.3.25}$$

其中权因子等于每次实验观测值误差平方的倒数

$$w_i = \frac{1}{(\Delta X_i)^2}.$$

当 H_0 为真, 并且每个实验中观测值服从正态分布时, $\hat{\mu}$ 和 $\hat{\sigma}$ 是总体参数的极大似然估计 (见式 (8.2.2) 和式 (8.2.3)).

若 H_0 为真, 变量 X^2(定义为各测量值 X_i 对加权平均 $\overline{X}$ 的离差平方的加权求和) 近似地服从 $\chi^2(N-1)$ 分布

$$X^2 \equiv \sum_{i=1}^{N} w_i (X_i - \overline{X})^2 = \sum_{i=1}^{N} \frac{(X_i - \overline{X})^2}{(\Delta X_i)^2} \tag{12.3.26}$$

(与式 (12.3.23) 对照). 因此, X^2 可作为 N 个测量具有相同均值的零假设 H_0 的检验统计量. 当 H_0 不为真, X^2 的值倾向于偏大, 因此, 临界域在 $\chi^2(N-1)$ 分布的上侧 (见 12.4 节的讨论). 给定显著性水平, 利用式 (12.3.7) 可确定上侧检验的临界域 (式中的 $Z \equiv X^2, f(z)$ 为 X^2 的概率密度). 利用测量值 $X_i, \Delta X_i$ 从式 (12.3.26) 求得 X^2 的观测值 X^2_{obs}, 若 X^2_{obs} 落在临界域内, 则拒绝零假设; 否则, 可认为零假设 H_0 成立, 并由式 (12.3.24) 和式 (12.3.25) 求出总体均值 μ 和方差 σ^2 的最佳估计值.

如果在选定的显著性水平上, X^2_{obs} 落在临界域内, 这时应当检查各个实验测定量 $X_i, \Delta X_i$ 对式 (12.3.26) 所表示的 X^2_{obs} 的贡献. 有时 X^2_{obs} 过大的原因在于一、二个实验测量值与其他测量值偏离过远, 使得它们对 X^2_{obs} 的贡献很大. 如果这种现象没有合理的解释, 建议去掉这些偏离很大的测量值, 对其余实验的结果重新进行检验.

例 12.7 Ω^- 超子质量

五个实验测得 Ω^- 超子质量值分别是 (1673.0 ± 8.0)MeV, (1673.3 ± 1.0)MeV, (1671.8 ± 0.8)MeV, (1674.2 ± 1.6)MeV, (1671.9 ± 1.2)MeV. 问在显著性水平 5%上这些结果是否一致.

由式 (12.3.26) 可算出 $X^2_{\text{obs}} = 3.24$. 作上侧检验, 由附表 7 查得 $\chi^2(N-1) = \chi^2(4)$ 累积分布函数值等于 $1-\alpha = 0.95$ 对应的上侧临界域为 $X^2 > 9.488$. 可见, X^2_{obs} 落在接受域内, 即在显著性 $\alpha = 0.05$ 上五个实验的结果是一致的.

由式 (12.3.24) 和式 (12.3.25) 确定 Ω^- 超子质量及其误差的估计值为 (1672.8 ± 0.5)MeV. 可见, 五个实验结果合并后, 其误差小于任何单个实验测定值的误差.

12.4 拟合优度检验

到目前为止, 本章前面各节讨论的内容属于参数检验问题, 要求解决的任务是根据一组观测值来确定某个概率分布的某一参数是否与零假设规定的参数值相一致. 在参数检验中, 分布的形式已经确定, 有待检验的是参数的数值.

本章的其余各节将讨论**非参数的假设检验**. 这有两方面的问题：一是本节将要讨论的**拟合优度检验**; 二是随机变量的**独立性、一致性检验**, 后者将在 12.6 节和 12.7 节中介绍.

拟合优度检验问题可表述如下：随机变量 X 的概率分布 $F(x)$ (连续或离散函数) 为未知, $X_1, X_2, \cdots, X_n$ 为变量 X 的一组随机子样. 令 $F_0(x)$ 为某个给定的概率分布函数, 现要求根据观测值 $X_1, X_2, \cdots, X_n$ 来检验 X 的概率分布是否可用 $F_0(x)$ 表示, 即零假设

$$H_0 : F(x) = F_0(x) \tag{12.4.1}$$

是否成立. 因此, 拟合优度检验的对象是概率分布的函数形式.

对于离散随机变量 X, 零假设 (式 (12.4.1)) 相当于

$$H_0 : P(x = x_i) \equiv p_i = p_{i0}, \quad i = 1, 2, \cdots.$$

对于连续随机变量 X, 零假设 (式 (12.4.1)) 相当于

$$H_0 : f(x) = f_0(x).$$

在作拟合优度检验时, 如同参数检验一样, 需要确定一个概率密度已知的检验统计量. 当 H_0 为真, 可根据该统计量的概率密度函数的积分和给定的显著性水平 α 来确定临界域. 与参数检验不同的是, 备择假设 H_1 只可能是与 H_0 不同的所有假设, 而不可能以某种特定的分布作为备择假设, 因此, 在拟合优度检验中, 备择假

设一般不专门给出. 由于检验的势函数取决于 H_1 的分布 (见 12.1.2 节的讨论), 因此, 拟合优度检验中对势函数不加考虑.

广泛采用的拟合优度检验是似然比检验和皮尔逊 χ^2 检验, 它们对于大样问题是严格正确的, 其他情形下只是近似地正确. 对于小子样容量的情形, 我们将讨论小样问题适用的柯尔莫哥洛夫检验.

12.4.1 似然比检验

将随机变量 (总体)X 的值域划分为互不相容的 N 个子区间, 设 $X_1, \cdots, X_{n_t}$ 为随机变量 X 的容量为 n_t 的一组子样观测值①, 落在 N 个子区间的事例数 (子样观测值个数) 为 $n = (n_1, \cdots, n_N)$, 显然我们有

$$\sum_{i=1}^{N} n_i = n_t \ . \tag{12.4.2}$$

当 H_0 为真, $\boldsymbol{n} = (n_1, \cdots, n_N)$ 的期望值 $\boldsymbol{\nu} = (\nu_1, \cdots, \nu_N)$ 由零假设概率密度函数决定

$$\nu_i = n_t \int_{x_i^{\min}}^{x_i^{\max}} f_0(x) \mathrm{d}x, \tag{12.4.3}$$

其中 $x_i^{\max}$ 和 $x_i^{\min}$ 是 i 子区间的上、下界. 利用观测数据 $\boldsymbol{n}$ 及其期望值 $\boldsymbol{\nu}$ 可以构造零假设 H_0 的拟合优度检验统计量.

对于 n_t 为固定常数的情形, 在 N 个子区间出现事例数期望值为$\boldsymbol{\nu}=(\nu_1, \cdots, \nu_N)$ 而观测事例数为 $\boldsymbol{n}=(n_1, \cdots, n_N)$ 的联合概率密度 (似然函数) 服从多项分布 (见 4.2 节)

$$L(\boldsymbol{n}|\boldsymbol{\nu}) = n_t! \prod_{i=1}^{N} \frac{1}{n_i!} \left(\frac{\nu_i}{n_t} \right)^{n_i}. \tag{12.4.4}$$

定义似然比

$$\lambda_M = \frac{L(\boldsymbol{n}|\boldsymbol{\nu})}{L(\boldsymbol{n}|\boldsymbol{n})} = \prod_{i=1}^{N} \left(\frac{\nu_i}{n_i} \right)^{n_i}, \tag{12.4.5}$$

当零假设 H_0 为真, 在子样容量 n_t 很大的极限情形下, 统计量

$$\chi_M^2 = -2 \ln \lambda_M = 2 \sum_{i=1}^{N} n_i \ln \frac{n_i}{\nu_i} \tag{12.4.6}$$

服从自由度 $N-1$ 的 χ^2 分布[35,57]. 自由度比子区间数 N 减少 1 是因为存在一个约束条件式 (12.4.2). 如果待检验的零假设 $H_0 : f(x) = f_0(x)$ 的概率密度函数中

① 注：为叙述方便, 本小节中子样容量记为 n_t, 其他小节中则记为 n

包含 k 个待估计参数 $\boldsymbol{\theta}=(\theta_1,\cdots,\theta_k)$, 它们用极大似然法求得其估计值 $\hat{\boldsymbol{\theta}}$, 代入式 (12.4.3) 求得 $\boldsymbol{\nu}=(\nu_1,\cdots,\nu_N)$ 的极大似然估计

$$\hat{\nu}_i=n_t\int_{x_i^{\min}}^{x_i^{\max}}f_0(x,\hat{\boldsymbol{\theta}})\mathrm{d}x. \tag{12.4.7}$$

当式 (12.4.6) 中的 ν_i 用极大似然估计 $\hat{\nu}_i$ 代替, 则统计量 χ_M^2 服从 $\chi^2(N-k-1)$ 分布 [35,57].

对于子样容量 n_t 为泊松变量 (期望值为 ν_t) 的情形, 落在 N 个子区间的事例数 (子样观测值个数) 为 $\boldsymbol{n}=(n_1,\cdots,n_N)$, 当 H_0 为真, 其期望值 $\boldsymbol{\nu}=(\nu_1,\cdots,\nu_N)$ 由零假设概率密度函数决定

$$\nu_i=\nu_t\int_{x_i^{\min}}^{x_i^{\max}}f_0(x)\mathrm{d}x, \tag{12.4.8}$$

那么, 在 N 个子区间出现事例数期望值为 $\boldsymbol{\nu}=(\nu_1,\cdots,\nu_N)$ 而观测事例数为 $\boldsymbol{n}=(n_1,\cdots,n_N)$ 的概率分布是泊松分布与多项分布的乘积[57]

$$L(\boldsymbol{n}|\boldsymbol{\nu})=\frac{\nu_t^{n_t}\mathrm{e}^{-\nu_t}}{n_t!}n_t!\prod_{i=1}^{N}\frac{1}{n_i!}\left(\frac{\nu_i}{\nu_t}\right)^{n_i}, \tag{12.4.9}$$

其中, $\nu_t=\sum\limits_{i=1}^{N}\nu_i$, $n_t=\sum\limits_{i=1}^{N}n_i$, 代入式 (12.4.9) 得

$$L(\boldsymbol{n}|\boldsymbol{\nu})=\prod_{i=1}^{N}\frac{\nu_i^{n_i}}{n_i!}\mathrm{e}^{-\nu_i}. \tag{12.4.10}$$

由该式表示的联合概率密度可以看到, 现在的情况相当于各子区间中的事例数 n_i 是相互独立的期望值为 ν_i 的泊松变量. 这一点在 4.4 节中已经提到: 当 N 个变量 $n_1,\cdots,n_N$ 服从多项分布, 而 $n_t=\sum\limits_{i=1}^{N}n_i$ 服从泊松分布时, 这 N 个变量是相互独立的泊松变量.

类似于式 (12.4.5) 和式 (12.4.6), 定义似然比 λ_P 和统计量 χ_P^2

$$\lambda_P=\frac{L(\boldsymbol{n}|\boldsymbol{\nu})}{L(\boldsymbol{n}|\boldsymbol{n})}=\mathrm{e}^{n_t-\nu_t}\prod_{i=1}^{N}\left(\frac{\nu_i}{n_i}\right)^{n_i}, \tag{12.4.11}$$

$$\chi_P^2=-2\ln\lambda_P=2\sum_{i=1}^{N}\left(n_i\ln\frac{n_i}{\nu_i}+\nu_i-n_i\right). \tag{12.4.12}$$

当零假设 H_0 为真, 在子样容量 n_t 很大的极限情形下, 统计量 χ_P^2 服从自由度 N 的 χ^2 分布[35,57]. 如果待检验的零假设 $H_0: f(x)=f_0(x)$ 的概率密度函数中包含 k

个待估计参数 $\boldsymbol{\theta}=(\theta_1,\cdots,\theta_k)$, 它们用极大似然法求得其估计值 $\hat{\boldsymbol{\theta}}$, 代入式 (12.4.8) 求得 $\boldsymbol{\nu}=(\nu_1,\cdots,\nu_N)$ 的极大似然估计

$$\hat{\nu}_i=\nu_t\int_{x_i^{\min}}^{x_i^{\max}}f_0(x,\hat{\boldsymbol{\theta}})\mathrm{d}x. \tag{12.4.13}$$

当式 (12.4.12) 中的 ν_i 用极大似然估计 $\hat{\nu}_i$ 代替, 则统计量 χ_P^2 服从 $\chi^2(N-k)$ 分布 [35,57].

这样, 就可以用检验统计量 χ_M^2 和 χ_P^2 来进行拟合优度检验. 给定显著性水平 α, χ^2 检验的上侧临界域的临界值由

$$\alpha=\int_{\chi_\alpha^2}^{\infty}f(u;v)\mathrm{d}u \tag{12.4.14}$$

决定, 其中 $f(u;v)$ 是自由度 v 的 χ^2 分布概率密度函数. 当 $\chi_M^2>\chi_\alpha^2$ 或 $\chi_P^2>\chi_\alpha^2$, 则在显著性水平 α 上拒绝零假设 H_0.

12.4.2　皮尔逊 χ^2 检验

当随机变量 X 为连续型, 将变量 X 的值域划分为互不相容的 N 个子区间 (如一维直方图中 N 个互不重叠的连续子区间), 第 i 子区间的变量值在 $x_{i-1}\to x_i$ 之间, 落入 i 子区间的事例数记为 n_i(观测频数), 事例总数为 n (注意：与 12.4.1 节的记号 n_t 不同), 则显然有 $\displaystyle\sum_{i=1}^{N}n_i=n$.

如果 H_0 成立, 第 i 子区间的理论频数为 np_{0i}, p_{0_i} 是事例落入 i 子区间的概率

$$p_{0_i}=\int_{x_{i-1}}^{x_i}f_0(x)\mathrm{d}x,\quad \sum_{i=1}^{N}p_{0_i}=1. \tag{12.4.15}$$

这相当于要求观测频数之和与理论频数之和相等

$$\sum_{i=1}^{N}n_i=n=\sum_{i=1}^{N}np_{0i}.$$

定义统计量

$$X^2\equiv\sum_{i=1}^{N}\frac{(n_i-np_{0i})^2}{np_{0i}}\equiv\frac{1}{n}\sum_{i=1}^{N}\frac{n_i^2}{p_{0i}}-n. \tag{12.4.16}$$

我们知道, 频率是概率分布的反映, 当子样容量 n 越来越大, 如果 H_0 为真, 频率 n_i/n 与概率 p_{0i} 之间的差别越来越小, 因此, 式 (12.4.16) 定义的 X^2 可以衡量子样观测值与总体分布 $f_0(x)$ 的差异程度, 于是 X^2 可考虑作为检验式 (12.4.1) H_0 假设的检验统计量.

当随机变量 X 为离散型, 零假设相当于分布律为

$$H_0 : P\left(X = x_i\right) \equiv p_i = p_{i0}, \quad i = 1, 2, \cdots, N.$$

令 n_i 为 X 取值 x_i 的观测事例数, 事例总数为 $n = \sum_{i=1}^{N} n_i$, 于是式 (12.4.16) 定义的 X^2 同样是适用的检验统计量.

为了确定临界域, 必须知道统计量 X^2 的分布, 定性地我们可作如下考虑: 在每一子区间中的观测频数 n_i 可考虑为泊松变量, 当 H_0 为真, 其均值和方差均为 np_{0i}. 当 np_{0i} 足够大, 泊松变量近似于均值和方差等于 np_{0i} 的正态变量 $N(np_{0i}, np_{0i})$, 因此

$$Y_i = \frac{n_i - np_{0i}}{\sqrt{np_{0i}}} \sim N(0,1), \quad i = 1, 2, \cdots, N. \tag{12.4.17}$$

X^2 正是这 N 个标准正态变量的平方和, 由于条件 (式 (12.4.15)) 的存在, 这 N 个变量中只有 $N-1$ 个是独立的. 因此, X^2 近似地服从 $\chi^2(N-1)$ 分布 (见 4.14 节). 统计量的这一性质称为**皮尔逊定理**: 不论分布 $f_0(x)$ 是何种函数, 当 H_0 为真, 统计量 X^2 的渐近分布 (当 $n \to \infty$) 是自由度 $N-1$ 的 χ^2 分布. 该定理的严格证明见文献 [35].

从皮尔逊定理可知, 式 (12.4.16) 定义的检验统计量 X^2 渐近地服从 $\chi^2(N-1)$ 分布这一性质与随机变量的分布函数 $f(x)$ 的形式无关, 因此, 这种检验是分布自由的, 适用于任何总体分布的随机变量的拟合优度检验问题.

根据以上的讨论, 在同样条件下, 实验重复多次得到各自的 n 个观测值, 如 H_0 为真, 按式 (12.4.16) 计算的 X^2 近似于 $\chi^2(N-1)$ 分布, 特别 X^2 的期望值 $\approx N-1$, 方差 $\approx 2(N-1)$. 若 H_0 不为真, n_i 的期望值不等于 np_{0i}, 则 $(n_i - np_{0i})^2$ 比 H_0 为真时要大, 即式 (12.4.16) 确定的 X^2 比 H_0 为真时要大, 因此, 对 H_0 应采用上侧检验. 这一点可从统计量 X^2 的期望值计算得到理论上的依据.

当 n 个观测值被分配在 N 个互不相容的连续子区间中, 第 i 个子区间内观测数为 n_i 个, 设事例落入第 i 子区间的概率为 p_i, 按 4.2 节所述, 随机变量 $(n_1, n_2, \cdots, n_N)$ 服从参数 n 和 $\boldsymbol{p} = (p_1, p_2, \cdots, p_N)$ 的多项分布. 统计量 X^2 的期望值是

$$E(X^2) = \frac{1}{n} \sum_{i=1}^{N} \frac{1}{p_{0i}} E(n_i^2) - n.$$

又

$$E(n_i^2) = np_i(1-p_i) + n^2 p_i^2,$$

故得

$$E(X^2) = \sum_{i=1}^{N} \frac{p_i(1-p_i)}{p_{0i}} + n\left(\sum_{i=1}^{N} \frac{p_i^2}{p_{0i}} - 1\right).$$

当 H_0 为真, 即 $p_i = p_{0i}$, $i = 1,\ 2,\ \cdots,\ N$, 对任意子样容量 n 均有

$$E(X^2|H_0) = N - 1.$$

将 $E(X^2)$ 的表达式对 p_i 求导数并等于 0, 以求得 $E(X^2)$ 的极小值, 并考虑到约束条件

$$\sum_{i=1}^{N} p_i = 1,$$

利用拉格朗日乘子法求解, 可发现, 当 $n \to \infty$ 时, 其极小值对应于 $p_i = p_{0i}$, 即 $E(X^2|H_0) = \min\left[E(X^2)\right]$. 因此, 对任何不同于原假设 H_0 的假设 H_1, 渐近地总有

$$E(X^2|H_1) > E(X^2|H_0).$$

该结果表明, 皮尔逊 χ^2 检验的临界域应当在 χ^2 分布的上侧.

给定显著性水平 α, 利用 $\chi^2(N-1)$ 的累积分布可确定临界值 $\chi^2_\alpha(N-1)$

$$\alpha = \int_{x^2_\alpha(N-1)}^{\infty} f(y; N-1)\mathrm{d}y, \tag{12.4.18}$$

其中, $f(y; N-1)$ 是 $\chi^2(N-1)$ 的概率密度, 临界域是 $\chi^2(N-1)$ 大于 $\chi^2_\alpha(N-1)$ 的区域. 对于一个实验的实际观测值 $n_i, i = 1, 2, \cdots, N$, 按照式 (12.4.16) 算出 X^2 的观测值 X^2_{obs}, 若 X^2_{obs} 大于 $\chi^2_\alpha(N-1)$, 则在显著性水平 α 上拒绝零假设; 当 $X^2_{\text{obs}} < \chi^2_\alpha(N-1)$, 则可认为总体的分布函数可由原假设 H_0 规定的 $f_0(x)$ 所描述.

由前面的讨论可见, 皮尔逊 χ^2 检验依赖于变量

$$Y_i = \frac{n_i - np_{0i}}{\sqrt{np_{0i}}}$$

渐近地服从标准正态分布这一假定 (见式 (12.4.17)), 这就要求落在每一子区间中的事例观测频数 n_i 充分大; 但另一方面, 将 n_i 个观测值归并到同一子区间中, 相当于用随机变量 X 在该子区间中的平均值代替 n_i 个不同的观测值, 这必定导致数据信息的某种损失. 这两者是相互矛盾的. 在实际选择子区间的大小时, 一般遵循的原则是, 使每个子区间内的理论频数 $\geqslant 5$ 的条件下, 子区间数目以较多为宜. 当自由度不太小 (如 $N-1 \geqslant 6$, N 子区间数), 在一个或两个子区间中理论频数甚至可以小于 5.

子区间的划分有两种方法. 一种是等宽度划分, 即各子区间宽度相等. 当随机变量 X 的取值域为 $(-\infty,\ \infty)$, 一般在两端概率密度极低, 故两端的两个子区间取为 $(-\infty, x_1)$ 和 (x_{N-1}, ∞). 这种划分方法的好处是十分简单, 采用得较为广泛. 另一种是等概率划分方法, 即在零假设 H_0 成立的条件下, 使每个子区间内的理论概

率相等 $p_1 = p_2 = \cdots = p_N$. 当子样容量 n 充分大, 在等概率划分的条件下, 可以确定出子区间数目 N 的一个最佳值, 使得皮尔逊 χ^2 检验的近似势函数达到极大[1]. 因此一般说来, 等概率方法在理论上较为优越, 但它的计算较为繁杂, 实际上使用不多.

在皮尔逊 χ^2 检验中, 差值 $n_i - np_{0i}$ 的符号对检验统计量 X^2 没有影响, 这表示该方法对数据所提供的信息尚没有充分地加以利用, 相应地, 检验的有效性会受到一定影响. 因此, 即使根据 X^2_{obs} 的值已接受 H_0 零假设的情况, 最好也检查一下观测频数与理论频数是否存在系统的偏差, 即 $n_i - np_{0i}$ 是否普遍地取相同的符号. 事实上, 即使皮尔逊 χ^2 检验中, X^2_{obs} 落入了接受域, 但零假设仍有可能被 $n_i - np_{0i}(i = 1, \cdots, N)$ 符号的一致性所排除, 这时游程检验可作为皮尔逊 χ^2 检验的必要补充 (见 12.7.3 节)

12.4.3 最小二乘、极大似然估计中的皮尔逊 χ^2 检验

12.4.2 节的讨论中我们假定了零假设 $H_0 : f(x) = f_0(x)$ 完全规定了 $f_0(x)$, 即 $f_0(x)$ 中不含未知参数. 但在许多实际问题中, 零假设 H_0 往往只给定了 $f_0(x)$ 的函数形式, 但含有若干个未知参数. 例如, 要求检验总体是否服从正态分布 $N(\mu, \sigma^2)$, 其中 μ 和 σ^2 均为未知. 这时式 (12.4.16) 检验统计量 X^2 的表达式中 p_{0i} 值无法求得, 因此, 12.4.2 节描述的检验方法不可能原封不动地采用, 而需要作某种修改. 一种自然的做法是利用这些未知参数的估计量来代替未知参数.

可以利用最小二乘估计量来代替未知参数. 当

$$Y_i = (n_i - np_{0i})/\sqrt{np_{0i}}$$

近似地服从标准正态分布时, 皮尔逊 χ^2 检验的统计量 X^2 的表达式与 9.4.3 节中 Q^2 的表达式一致 (见式 (9.4.14)). 在 Q^2 表达式中, N 个测量值是相互独立的, Q^2 是 $\chi^2(N)$ 变量; 而现在 X^2 表达式中 N 个测量值 n_i 只有 $N-1$ 个是独立的, 因而 $X^2 \sim \chi^2(N-1)$, 这一点已经在 12.4.2 节阐明. 如果总体中存在 L 个未知参数, 可通过求 Q^2 的极小值 $Q^2_{\min}$ 得到这 L 个参数的最小二乘估计, 代入 Q^2 的表式得到 $Q^2_{\min}$, 而 $Q^2_{\min} \sim \chi^2(N-L)$(见 9.4.3 节). 在现在的情形下, 可通过求 X^2 的极小值 $X^2_{\min}$ 得到 L 个参数的最小二乘估计, 而 $X^2_{\min}$ 近似地服从 $\chi^2(N-1-L)$ 分布. 因此, 对于零假设中含有 L 个独立参数的情形, 当给定显著性水平 α 后, 临界域由 $\chi^2(N-L-1)$ 的概率密度的积分 (积分限为 $\chi^2_\alpha(N-1-L)$ 和 ∞) 决定. 当 $X^2_{\min}$ 的实际观测值落入临界域中, 则在水平 α 上拒绝零假设; 反之则接受零假设, 零假设的未知参数由最小二乘估计确定. 需要指出的是, 以上结论仅适用于参数的线性模型; 对于非线性的最小二乘估计, 上述方法只是一种近似.

也可以利用极大似然法来对未知参数进行估计, 并由此求得式 (12.4.16) 中 p_{0i} 的值. 可以考虑两种可能性：①在作参数估计时, 观测数据划分为 N 个子区间, L 个未知参数用 8.7 节描述的方法进行估计. 这样求出的 p_{0i} 的估计值 $\hat{p}_{0i}$ 代入式 (12.4.16) 求出统计量 X^2, 这时, X^2 服从 $\chi^2(N-1-L)$ 或 $\chi^2(N-1)$ 分布, 如同用最小二乘法估计参数的情形相同. ②L 个未知参数用一般的极大似然法进行估计, 求出的 $\hat{p}_{0i}$ 代入式 (12.4.16) 得到统计量 X^2, 这时, X^2 的分布不再是简单的 χ^2 分布, 而是位于 $\chi^2(N-1)$ 和 $\chi^2(N-1-L)$ 之间[1]. 但当 N 足够大时, $\chi^2(N-1)$ 和 $\chi^2(N-1-L)$ 的差别不显著. 当给定显著性水平 α 时, X^2 的临界域可从 $\chi^2(N-1)$ 的概率密度找出. 对于 N 不够大的场合, 为可靠起见, 可要求 X^2_{obs} 超过 $\chi^2(N-1)$ 和 $\chi^2(N-1-L)$ 求出的临界值中的小者即否定零假设 H_0. 由于 $\chi^2_\alpha(N-1-L) < \chi^2_\alpha(N-1)$, 这等效于临界域为 $\left(\chi^2_\alpha(N-1-L), \infty\right)$.

12.4.4 拟合优度的一般 χ^2 检验

在 12.4.2 节 ~12.4.3 节的讨论中, 为了构造检验统计量, 我们将随机变量的值域划分为 N 个子区间, 并要求这 N 个子区间中理论频数总和与观测频数总和相等, 即要求 $\sum\limits_{i=1}^{N} p_{0i} = 1$. 这一归一化要求使得检验统计量 X^2 的自由度减少了 1.

但在某些问题中, 这种归一化的要求是不合理或不必要的. 例如, 随机变量 X 的值域为 $(-\infty, \infty)$, 但在实验中, 由于测量仪器的限制, 只可能测到有限的区间, 因此, 合理的做法是在有限的区间内进行拟合检验. 这时, 归一化要求便是不合理的, 检验的模型预期的理论频数总和不要求与观测频数总和相等. 因此, 一般的待检验的零假设可表示为

$$H_0 : f_1 = f_{01}, \quad f_2 = f_{02}, \ \cdots, \ f_N = f_{0N}, \tag{12.4.19}$$

其中 $f_1, f_2, \cdots, f_N$ 为理论频数. 检验统计量现在是

$$X^2 = \sum_{i=1}^{N} \frac{(n_i - f_{0i})^2}{f_{0i}}. \tag{12.4.20}$$

当 H_0 为真且所有的观测频数 n_i 充分大 (如大于 5), 式 (12.4.20) 右边求和号中的每一项近似地服从 $N(0,\ 1)$ 分布, X^2 近似地为 $\chi^2(N)$ 变量. 如果零假设 H_0 包含 L 个独立的未知参数, 可用一般的估计方法求得参数的估计值后再进行拟合优度检验, 将估计值 $\widehat{f_{0i}}$ 代替 f_{0i} 求得的 X^2 表达式近似地服从 $\chi^2(N-L)$ 分布. 这样, 当给定了显著性 α 后, 容易求出相应的临界域.

对于更为复杂的问题, 例如, 各测量值真值之间不相独立而是存在着相互关联, N 个测量值真值与不可直接测量的 L 个未知参数之间存在着 K 个独立约束方程

的情形 (见 9.8 节的讨论), 拟合优度检验可选择下述检验统计量:

$$X^2 = (Y - f_0)^{\mathrm{T}} \underset{\sim}{V}^{-1}(Y)(Y - f_0), \tag{12.4.21}$$

式中 $Y = (Y_1, Y_2, \cdots, Y_N)$ 是观测值向量, $\underset{\sim}{V}(Y)$ 是它的协方差矩阵, f_0 是零假设 H_0 对 Y 的拟合量向量. 当用最小二乘法作参数估计时, 待估计的未知量共 $N + L$ 个 (N 个测量值真值 f_{0i}, $i = 1, 2, \cdots, N$, L 个未知参数), 利用 K 个约束方程消去 K 个未知量, 独立的未知量为 $N + L - K$ 个. 将它们的最小二乘估计代入式 (12.4.21), 求得 $X^2_{\min}$, 它服从 $\chi^2(\nu)$ 分布, 其自由度为测量个数 N 减去独立的未知量个数

$$\nu = N - (N + L - K) = K - L.$$

给定显著性水平 α, χ^2 检验的临界域的临界值 χ^2_α 由

$$\alpha = \int_{\chi^2_\alpha}^{\infty} f(u; v)\mathrm{d}u = 1 - F(u = \chi^2_\alpha; v) \tag{12.4.22}$$

给定, 其中 $v = K - L$, $f(u; v)$ 是 $\chi^2(v)$ 的概率密度, $F(u; v)$ 是其累积分布函数. 如果由观测值算得的 X^2_{obs} 超过 χ^2_α, 则在显著性水平 α 上拒绝假设 H_0.

物理学家普遍采用的一种做法是将实际观测值 X^2_{obs} 对应的积分概率 P_{x^2} 作为表征拟合优度的特征量 (见 9.4.4 节)

$$P_{x^2} = \int_{X^2_{\mathrm{obs}}}^{\infty} f(u; v)\mathrm{d}u = 1 - F(u = X^2_{\mathrm{obs}}; v). \tag{12.4.23}$$

容易看到下面两式是等价的:

$$X^2_{\mathrm{obs}} > \chi^2_\alpha, \quad P_{x^2} < \alpha. \tag{12.4.24}$$

因此, 后一个式子同样可作为在显著性水平 α 上拒绝 H_0 的判别式.

应当强调, 当 $X^2_{\mathrm{obs}} > \chi^2_\alpha (P_{x^2} < \alpha)$ 时, 表示拟合很差, 有充分的理由在显著性水平 α 上拒绝零假设. 相反, $X^2_{\mathrm{obs}} < \chi^2_\alpha (P_{x^2} > \alpha)$ 时却不能肯定地接受假设 H_0. 因为正如我们在 12.4.2 节中所指出的, 皮尔逊 χ^2 检验对差值 $Y_i - f_{0i}$ 的符号信息没有加以利用. 因此, 比较恰当的说法是, $X^2_{\mathrm{obs}} < \chi^2_\alpha (P_{x^2} > \alpha)$ 时, 在显著性水平 α 上不能排除假设 H_0. 如果一个物理学家仅仅从皮尔逊检验得出接受某个零假设的结论, 可能不是仅仅从纯粹统计学的观点出发, 而是加上了物理上的考虑.

利用检验统计量的观测值 X^2_{obs} 到 ∞ 的积分概率 P_{x^2} 作为检验 H_0 成立与否的判别量 (式 (12.4.24)), 可推广到更一般的场合. 令检验零假设 H_0 的检验统计量为 Q, H_0 为真时, Q 的概率密度和累积分布分别为 $f(Q)$ 和 $F(Q)$, Q 的观测值为 Q_{obs}, Q_{obs} 的积分概率定义为

$$P_{\mathrm{obs}} = \int_{Q_{\mathrm{obs}}}^{\infty} f(Q)\mathrm{d}Q = 1 - F(Q_{\mathrm{obs}}). \tag{12.4.25}$$

给定显著性水平 α, 检验可分下列三种情况.

(1) 上侧检验：Q 的临界值 Q_α 由下式决定：

$$\alpha = \int_{Q_\alpha}^{\infty} f(Q)\mathrm{d}Q = 1 - F(Q_\alpha).$$

于是当

$$P_{\text{obs}} < \alpha \quad (\text{即} Q_{\text{obs}} > Q_\alpha) \tag{12.4.26}$$

时, 拒绝 H_0.

(2) 下侧检验：Q_α 为临界值, 则满足

$$\alpha = \int_{-\infty}^{Q_\alpha} f(Q)\mathrm{d}Q = F(Q_\alpha).$$

当

$$P_{\text{obs}} > 1 - \alpha \quad (\text{即} Q_{\text{obs}} < Q_\alpha) \tag{12.4.27}$$

时, 拒绝 H_0.

(3) 双侧检验：Q 的上、下临界值为 Q_u, Q_d,

$$\frac{\alpha}{2} = \int_{-\infty}^{Q_d} f(Q)\mathrm{d}Q = F(Q_d) = \int_{Q_u}^{\infty} f(Q)\mathrm{d}Q = 1 - F(Q_u).$$

拒绝 H_0 的条件为

$$P_{\text{obs}} > 1 - \frac{\alpha}{2} \quad (\text{即} Q_{\text{obs}} < Q_d) \tag{12.4.28}$$

或

$$P_{\text{obs}} < \frac{\alpha}{2} \quad (\text{即} Q_{\text{obs}} > Q_u).$$

例 12.8　$\mathrm{e}^+\mathrm{e}^- \to \mu^+\mu^-$ *反应的角分布* (2)

设正、负电子在同一直线上相向 “碰撞”, 而且它们有相等的能量, 该反应的可能末态之一是产生一对 $\mu^+\mu^-$ 粒子, 能量相同而方向相反. 令 e^+ 与 μ^+ 之间的夹角为 ϑ, φ(ϑ 是极角, φ 是方位角). 根据量子电动力学 (下面以 QED 表示) 理论, 当 $\mathrm{e}^+\mathrm{e}^-$ 的能量不高时, 反应的角分布为

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\mathrm{d}\sigma}{\mathrm{d}\varphi \mathrm{d}\cos\vartheta} = \frac{\alpha^2}{16E^2}(1+\cos^2\vartheta),$$

其中 E 是入射 e^+(或 e^-) 的能量, 该分布对 $\cos\vartheta = 0$ 为对称. 当能量 E 高时, 除了 QED 考虑的电磁作用外, 还需考虑弱作用项的贡献, 因此, 弱电统一理论 (以下用 EW 表示) 预期角分布为

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\alpha^2}{16E^2}\left[(1+a_1)(1+\cos^2\vartheta) + a_2\cos\vartheta\right],$$

其中 a_1, a_2 是与 E 有关的量. 上式第二项对于 $\cos\vartheta = 0$ 不对称, 而且 a_1, a_2 的值使得当 E 增大时不对称性也增大. 通过对 $e^+e^- \to \mu^+\mu^-$ 反应角分布的测量, 可对弱作用进行研究.

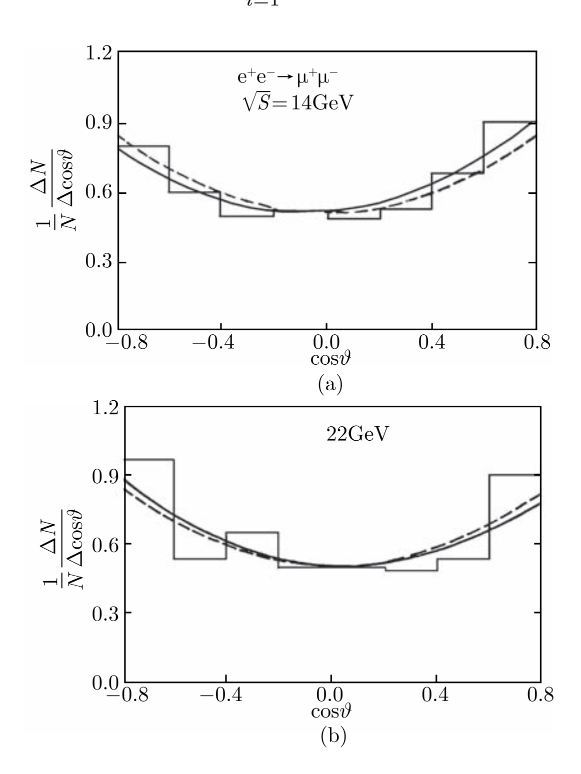
在 $E = 7\text{GeV}$, 11GeV, 17.3GeV 三种能量下, 测量了 $e^+e^- \to \mu^+\mu^-$ 事例的角分布, 用直方图的形式示于图 12.7 (a), (b), (c), 其中还给出了 QED(虚线) 和 EW(实线) 理论的预期值, N 表示总的事例数, $\Delta N/\Delta\cos\vartheta$ 表示落在 $\Delta\cos\vartheta$ 间隔内的事例数, 因此, $\dfrac{1}{N}\dfrac{\Delta N}{\Delta\cos\vartheta}$ 为归一分布. 具体测量数据见表 12.3, 其中理论值 f_i 表示 EW 理论对落在子区间 $\Delta\cos\vartheta_i$ 中事例数的预期, 并且有

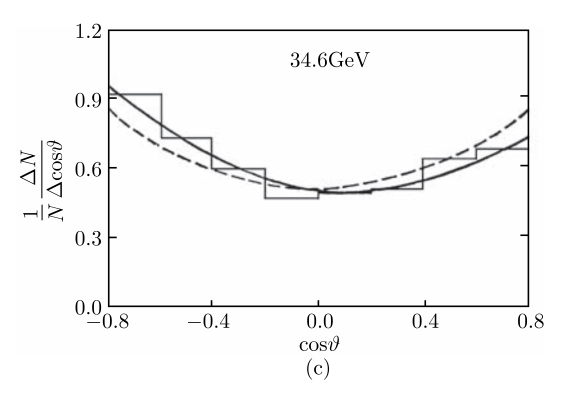
$$\sum_{i=1}^{8} n_i = \sum_{i=1}^{8} f_i,$$

n_i 为相应的实测值, 见表 12.3.

我们用皮尔逊 χ^2 方法来检验理论与实测值的一致性. 由于共有 8 个子区间, 并且有关系式 $\sum\limits_{i=1}^{8} n_i = \sum\limits_{i=1}^{8} f_i$, 因此, 量

$$X^2 = \sum_{i=1}^{8} \frac{(n_i - f_i)^2}{f_i}$$





(c)

图 12.7　三种能量下的 $e^+e^- \to \mu^+\mu^-$ 角分布

(a) $E = 7$ GeV; (b) $E =11$ GeV; (c) $E = 17.3$ GeV

直方图 —— 测量值; 虚线 —— 量子电动力学预期值; 实线 —— 弱电统一理论预期值

表 12.3　$e^+e^-\to\mu^+\mu^-$ 反应角分布的测量数据

$\Delta\cos\vartheta$ \ E/GeV	7		11		17.3	
	实验 n_i	理论 f_i	实验 n_i	理论 f_i	实验 n_i	理论 f_i
$-0.8/-0.6$	74	66.5	71	60.4	426	404
$-0.6/-0.4$	55	56.0	39	50.1	336	334
$-0.4/-0.2$	46	50.0	48	43.0	277	280
$-0.2/0$	48	47.8	36	38.6	219	245
$0/0.2$	45	49.1	36	38.0	227	232
$0.2/0.4$	49	54.1	36	40.8	234	243
$0.4/0.6$	63	63.3	39	46.5	295	295
$0.6/0.8$	82	75.2	67	54.6	313	313
总事例数	462		372		2327	

服从 $\chi^2(7)$ 分布. 对于三种能量值 $E = 7$GeV, 11GeV, 17.3GeV, 观测值 $X^2_{\rm obs}$=2.62, 9.77, 5.95. 若取显著性水平 $\alpha = 0.10$, 则临界值为 $\chi^2_\alpha(7)$=12.02, 故三种能量值的观测数据都满足

$$X^2_{\rm obs} < \chi^2_{\alpha=0.10}(7),$$

即没有理由认为实验测量的结果与弱电统一理论不一致.

如果将实验数据与 QED 的预期相比较, 由图可以直观地看到对于 $E = 7$GeV, 11GeV 的情形, 两者相当好地一致, 皮尔逊检验得出同样的结论; 而当 E=17.3GeV 时, 两者相差甚大, 事实上, 通过计算可知, $X^2_{\rm obs} \approx 48$, 在一般采用的显著性水平上 (α=0.10, 0.05), 都排除了实测值与 QED 理论一致的可能性.

以上结果反映出, 在能量 E 较低时, 由于弱作用贡献的不对称项比较小, EW 与 QED 理论对角分布的预期值差别不大, 实验测量又不够精确, 因而无法在两种

理论中作出鉴别; 但当能量升高时, 弱作用项的贡献足够大, 实验数据只能用弱电统一理论才能解释.

例 12.9 $e^+e^- \to \mu^+\mu^-$ 反应总截面随能量的分布

QED 理论预期, $e^+e^- \to \mu^+\mu^-$ 反应总截面可表示为

$$\sigma_{\mu\mu} = \frac{21.714\text{nb}}{E^2},$$

其中 nb(10^{-33}cm^2) 是反应截面单位. 在不同能量处观测到的事例数 n_i 如表 12.4 所示. 表中 L_i 表示实验测量时在能量 E_i 处的积分亮度, 理论事例数 f_i 与 L_i 有简单的关系

$$f_i = AL_i\sigma_{\mu\mu},$$

A 是由实验装置决定的一个常数.

表 12.4 $e^+e^-\to\mu^+\mu^-$ 总截面能量依赖的测量数据

E_i/GeV	L_i/nb^{-1}	$\sigma_{\mu\mu}$/nb	f_i(理论)	n_i(实验)
6	97.7	0.603	26.7	31
6.5	53	0.514	11.6	13
8.5	60	0.301	7.7	16
11	51.5	0.179	3.9	3
13.7	483	0.116	23.8	18
15	623	0.0965	25.6	28
15.3	2015	0.0928	84.5	91
15.8	238	0.0870	8.8	7
16.7	1785	0.0779	64	66
17.24	2257	0.0731	76.3	58
17.72	3084	0.0692	99.2	110
18.2	1854	0.0656	56.8	49

现在来检验实验测量与理论的一致性. 由于测量是在若干个能量点进行的, 显然不可能覆盖 E 的全部区域, 因此, 理论值与观测值之间的归一既不可能也不必要, 应当使用本节描述的拟合优度的一般 χ^2 检验. QED 对 $e^+e^- \to \mu^+\mu^-$ 反应的截面公式作为零假设, 那么在实验给定的积分亮度下, 对反应事例数的理论预期值如表 12.4 中的 f_i 所表示. 根据式 (12.4.20), 量

$$X^2 = \sum_{i=1}^{12} \frac{(n_i - f_i)^2}{f_i}$$

近似地服从 $\chi^2(12)$ 分布. 由于自由度 $N = 12$ 相当大, 虽然 E=11GeV 的理论事例数 (3.9 个) 小于 5, 按照 12.4.2 节的讨论, 该近似仍然适用. 取显著性水平 α= 0.05,

确定出临界值 $\chi^2_{\alpha=0.05}(12) = 21.03$. 实验观测值 $X^2_{\rm obs}$ 计算结果为

$$X^2_{\rm obs} = 19.22 < 21.03,$$

所以在显著性 α=0.05 水平上没有理由认为实验与理论不一致.

利用事例数与积分亮度的简单关系容易算出 $\sigma_{\mu\mu}$ 的实验测定值, 画在图 12.8 中. 可以看到, 理论与测量的最大偏差出现在 $E \equiv \sqrt{S}/2$= 8.5GeV 处, 在此情形下, 对该点的实验数据作仔细的检查是必要的.

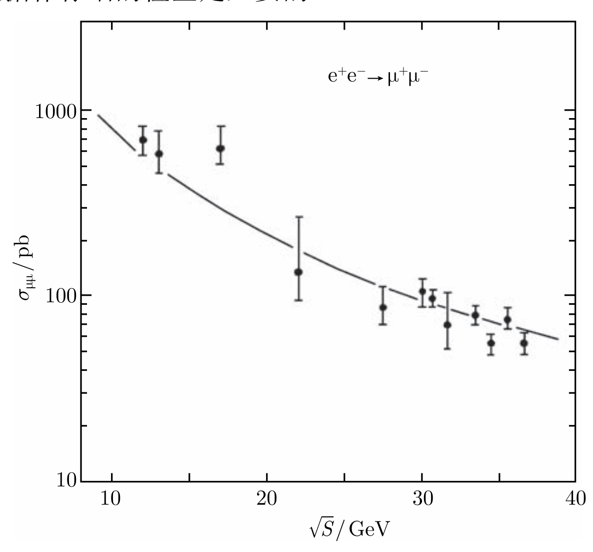


图 12.8　$\rm e^+e^- \to \mu^+\mu^-$ 反应截面的理论预期和测量值

例 12.10　V^0 事例的运动学分析 (2)

中性奇异粒子 Λ^0 和 $\rm K^0_{\it s}$ 的衰变方式分别为

$$\Lambda^0 \to \rm p^+ + \pi^-, \qquad K^0_{\it s} \to \pi^+ + \pi^-.$$

它们在气泡室中观测到的都是 V^0 形事例 (见例 9.12 的说明). 我们可以用拟合优度的皮尔逊 χ^2 检验来鉴别衰变的中性粒子是 Λ^0 还是 $\rm K^0_{\it s}$. 对每个 V^0 事例, 假定只存在两种假设

$$H_0 : \Lambda^0 \to \rm p^+ + \pi^-, \qquad H_1 : K^0_{\it s} \to \pi^+ + \pi^-.$$

每个粒子由三个运动学变量描述：动量 p, 飞行方向的极角 ϑ 和方位角 φ. 假定中性粒子 V^0 的 ϑ、φ 已从该粒子的产生点和衰变点确定, 衰变产生的两个带电粒子的动量和飞行方向 ϑ、φ 也根据气泡室的径迹分析测量到了, 于是测量值 η 是一个

8 个分量的向量, 唯一的未知参数 ξ 是中性粒子动量数值. 对于这 9 个运动学变量, 由于衰变前后必须服从能量、动量守恒, 故共存在 4 个独立的约束方程.

利用最小二乘法求得拟合值, $\boldsymbol{f}_0 = \widehat{\boldsymbol{\eta}}$, 代入式 (12.4.21) 求出 $X^2_{\min}$, $X^2_{\min}$ 将服从 $\chi^2(3)$ 分布. 当给定显著性水平 α, 即可求出临界域, 对零假设作出接受或拒绝的结论.

设一个特定的 V^0 事例有以下数据：V^0 粒子的产生点和衰变点坐标分别为 $(-44.40\pm0.14, -1.80\pm0.17, -16.20\pm0.26)$ 和 $(-28.80\pm0.15, -5.30\pm0.16, -16.00\pm0.26)$, 单位是厘米. 由这些数据可求得 V^0 粒子的飞行极角和方位角 ϑ、φ. 两个带电粒子的动量值, 极角 ϑ 和方位角 φ 如表 12.5 所列, 当带正电的粒子考虑为质子 p 和介子 π^+ 时, ϑ、φ 值有微小的差别, 这是由于物理上的考虑 (粒子径迹的多次散射修正). 所有测量值的误差互不相关.

表 12.5 一个 V° 事例的测量数据及其拟合值

		动量/(MeV/c)	极角 ϑ/弧度	方位角 φ/ 弧度
测量值	正粒子设为 p	1535±72	1.549±0.006	6.107±0.007
	正粒子设为 π^+	1479±60	1.552±0.006	6.111±0.006
	负粒子 π^-	378±18	1.668±0.016	5.768±0.014
H_0 为真时的拟合值	p	1564±72	1.549±0.006	6.106±0.007
	π^-	354±11	1.662±0.016	5.781±0.012
H_1 为真时的拟合值	π^+	1831±51	1.547±0.006	6.124±0.006
	π^-	381±18	1.689±0.016	5.719±0.011

给定显著性水平 $\alpha = 0.01$ 由附表 7 查得上侧临界值为 $\chi^2_{0.01}(3)$=11.345, 而观测值 $X^2_{\text{obs}}(H_0)$=3.6, $X^2_{\text{obs}}(H_1)$=26.7. 显然应当排除 H_1 而接受 H_0, 故这一事例为 $\Lambda^0 \to \mathrm{p} + \pi^-$ 衰变事例.

本例中, $X^2_{\text{obs}}(H_0)$ 和 $X^2_{\text{obs}}(H_1)$ 相差很大, 明显地落在临界值 $\chi^2_{0.01}(3)$=11.345 的两边, 这使得事例的鉴别十分容易. 但有的情况下事情并非这样截然分明. 例如, 两个假设都满足 $X^2_{\text{obs}} < \chi^2_\alpha$, 则在显著性水平 α 上这两种假设都可接受, 这个事例便是模棱两可的. 如果无法用其他测量作进一步的判别, 一般的做法是取 X^2_{obs} 较小的那个假设. 当两个假设都满足 $X^2_{\text{obs}} > \chi^2_\alpha$, 则该事例应当排除于这两个假设规定的事例之外.

12.4.5 柯尔莫哥洛夫检验

皮尔逊 χ^2 检验无疑是物理学家使用得最普遍的非参数检验方法. 但正如前面所指出的, 它只适用于子样容量大的场合, 而且由于将数据归并到 N 个子区间中, 导致某种程度的信息损失, 降低了检验的有效性. 本节所要阐述的柯尔莫哥洛夫 (Kolmogorov) 检验方法, 避免了对数据的分组划分, 因而更充分地利用了数据的

信息; 同时该方法对任何子样容量都适用, 故对小样问题明显地比皮尔逊 χ^2 检验优越. 但是应当注意, 柯尔莫哥洛夫检验仅适用于连续随机变量, 并且零假设给定的分布不能包含未知参数.

设 $x_1, x_2, \cdots, x_n$ 是随机变量 X 的顺序统计量 (顺序统计量的定义见 6.2 节) 的一组观测值, 即

$$x_1 \leqslant x_2 \leqslant x_3 \cdots \leqslant x_n.$$

子样分布函数 $S_n(x)$ 定义为 (见 6.1 节)

$$S_n(x) = \begin{cases} 0, & x < x_1, \\ \dfrac{k}{n}, & x_k \leqslant x < x_{k+1}, \quad k = 1, \cdots, n-1, \\ 1, & x_n \leqslant x. \end{cases} \tag{12.4.29}$$

$S_n(x)$ 是一上升的阶梯函数, 在每一观测值 $x_1, x_2, \cdots, x_n$ 处阶梯增高 $1/n$.

柯尔莫哥洛夫检验是利用子样分布函数 $S_n(x)$ 来检验随机变量 X 的分布函数 $F(x)$ 是否具有特定的形式 $F_0(x)$, 即零假设取为

$$H_0 : F(x) = F_0(x). \tag{12.4.30}$$

在 6.1 节中已经指出, 格利汶科定理证明了当 $n \to \infty$ 时, $S_n(x)$ 依概率 1 收敛于总体 X 的累积分布函数 $F(x)$. 可以预计, 当 H_0 为真, 即 $F(X) = F_0(x)$, 则在任一 x_i, $i = 1, \ldots, n$ 处, $S_n(x)$ 与 $F_0(x)$ 的值应当十分接近; 相反, 若在某些 x_i 处 $S_n(x)$ 与 $F_0(x)$ 的值差异很大, 则总体 x 的分布与给定分布 $F_0(x)$ 处处相符的可能性很小, 即原假设很可能不为真. 因此, $S_n(x)$ 与 $F_0(x)$ 的差值可以作为检验统计量.

我们定义如下三个随机变量来表征 $S_n(x)$ 与 $F_0(x)$ 的差值：

$$\begin{aligned} D_n^+ &= \max_{-\infty<x<+\infty} (S_n(x) - F_0(x)), \\ D_n^- &= \max_{-\infty<x<+\infty} (F_0(x) - S_n(x)), \\ D_n &= \max_{-\infty<x<+\infty} |S_n(x) - F_0(x)| = \max(D_n^+, D_n^-). \end{aligned} \tag{12.4.31}$$

当总体 X 的分布 $F(x)$ 为连续函数, 则 D_n 和 D_n^+ 的概率分布为 (D_n^+ 与 D_n^- 有相同的分布)[14]

$$P\left(D_n < z + \frac{1}{2n}\right) = \begin{cases} 0, & z \leqslant 0, \\ 1, & z \geqslant 1 - \dfrac{1}{2n}, \\ \displaystyle\int_{\frac{1}{2n}-z}^{\frac{1}{2n}+z} \int_{\frac{3}{2n}-z}^{\frac{3}{2n}+z} \cdots \int_{\frac{2n-1}{2n}-z}^{\frac{2n-1}{2n}+z} f(y_1, \cdots, y_n) \mathrm{d}y_1 \cdots \mathrm{d}y_n; & 0 < z < 1 - \dfrac{1}{2n}, \end{cases}$$

$$P(D_n^+ \leqslant z) = \begin{cases} 0, & z \leqslant 0, \\ 1, & z \geqslant 1, \\ \int_{1-z}^{1} \int_{\frac{n-1}{n}-z}^{y_n} \cdots \int_{\frac{2}{n}-z}^{y_3} \int_{\frac{1}{n}-z}^{y_2} f(y_1, \cdots, y_n) \mathrm{d}y_1 \cdots \mathrm{d}y_n; & 0 < z < 1. \end{cases}$$

式中,

$$f(y_1, \cdots, y_n) = \begin{cases} n!, & 0 < y_1 < \cdots < y_n < 1, \\ 0, & \text{其他}. \end{cases}$$

在子样容量 n 很大时的极限分布为

$$\lim_{n \to \infty} P\left(D_n \leqslant \frac{z}{\sqrt{n}}\right) = 1 - 2\sum_{i=1}^{\infty} (-1)^{i-1} \mathrm{e}^{-2i^2 z^2} \ (z > 0), \tag{12.4.32}$$

$$\lim_{n \to \infty} P\left(D_n^+ \leqslant \frac{z}{\sqrt{n}}\right) = 1 - \mathrm{e}^{-2z^2} \ (z > 0), \tag{12.4.33}$$

由式 (12.4.33) 立即有

$$\lim_{n \to \infty} \left\{4n(D_n^+)^2 \leqslant z\right\} = 1 - \mathrm{e}^{-\frac{z}{2}}.$$

与式 (4.14.2)$\chi^2(n)$ 的累积分布函数对比, 即知, 当 $n \to \infty$ 时

$$D_n^+ \sim \sqrt{\frac{\chi^2(2)}{4n}}. \tag{12.4.34}$$

可见, 统计量 D_n, D_n^+, D_n^- 的分布与总体分布 $F(x)$ 无关, 且这一结论对任意子样容量 n 皆为正确. 因此, 利用 D_n, D_n^+, D_n^- 作为检验统计量的柯尔莫哥洛夫检验是分布自由的, 适用于任何连续总体.

对于不同的备择假设, 需选用不同的检验统计量, 见表 12.6. 当 H_0 为真, $D_n(D_n^+,\ D_n^-)$ 接近于 0; 当 H_0 不为真而备择假设 H_1 为真, $D_n(D_n^+,\ D_n^-)$ 有增大的趋势. 因此, $D_n(D_n^+, D_n^-)$ 的临界域在其分布的上侧. 给定显著性水平 α, 其临界值 $D_{n,\alpha}^{(+,-)}$ 由

$$P\left\{D_n^{(+,-)} > D_{n,\alpha}^{(+,-)}\right\} \leqslant \alpha \tag{12.4.35}$$

给出. 对于五种不同的显著性水平 α 值, $D_n^{(+,-)}$ 的临界值列于附表 12, 表的最后一行给出了 n 趋于 ∞ 时的临界值 $D_{n,\alpha,n\to\infty}$, 并且总有

$$D_{n,\alpha} < D_{n,\alpha,n\to\infty}.$$

对于单侧备择假设的情形, 当 n 很大时, 还可利用式 (12.4.34) 来确定临界值

$$D_{n,\alpha}^{+,-} = \sqrt{\frac{\chi_\alpha^2(2)}{4n}}, \tag{12.4.36}$$

其中 $\chi^2_\alpha(2)$ 是 $\chi^2(2)$ 分布的上侧 α 分位数.

表 12.6　柯尔莫哥洛夫检验中的备择假设和检验统计量

原假设	备择假设	检验统计量	注
$H_0 : F(x) = F_0(x)$	$H_1 : F(x) \neq F_0(x)$	D_n	双侧备择假设
$H_0 : F(x) = F_0(x)$	$H_1 : F(x) > F_0(x)$	D_n^+	单侧备择假设
$H_0 : F(x) = F_0(x)$	$H_1 : F(x) < F_0(x)$	D_n^-	单侧备择假设

这样, 根据一组观测值 $x_1, \cdots, x_n$ 来检验零假设 H_0 就变得十分简单. 对于所需要的备择假设选择适当的统计量 D_n 或 D_n^+, D_n^-, 将观测值 $x_1, \cdots, x_n$ 代入式 (12.4.31) 得到 $D_n^{(+,-)}$ 的实际观测值 $D_{n,\text{obs}}^{(+,-)}$, 如果 $D_{n,\text{obs}}^{(+,-)} > D_{n,\alpha}^{(+,-)}$, 则在显著性水平 α 上拒绝 H_0 而接受备择假设; 反之, 则接受原假设 H_0.

对 $H_1 : F(x) \neq F_0(x)$ 的情形, 式 (12.4.35) 结合式 (12.4.31) 可重新表述为

$$P\left\{S_n(x) - D_{n,\alpha} \leqslant F_0(x) \leqslant S_n + D_{n,\alpha}\right\} \geqslant 1 - \alpha \quad (\text{对所有}x). \tag{12.4.37}$$

该式表示累积分布函数 $F_0(x)$ 大于 $S_n(x) - D_{n,\alpha}$ 而小于 $S_n(x) + D_{n,\alpha}$ 的概率大于 $1 - \alpha$, 即真的累积分布函数 $F_0(x)$ 落在 $[S_n(x) - D_{n,\alpha}, S_n(x) + D_{n,\alpha}]$ 区域内的置信概率大于 $1 - \alpha$. 这一关系可以用来估计为了使测量值能反映总体分布达到一定的精度, 需要多少次测量 (即子样容量 n). 例如, 要求在置信水平 90%上, 实验测量得到的分布反映真实分布 $F_0(x)$ 的精度好于 0.20. 因为 D_n 是子样分布与真分布 $F_0(x)$ 的差值, 它是测量值反映 $F_0(x)$ 的精度的直接表示, 所以上述要求相当于 $\alpha = 10\%$, $D_n \leqslant 0.20$. 从附表 12 可以查到, 为了使 $D_n \leqslant 0.20$, 子样容量必须满足 $n > 35$. 也就是说, 36 次测量构成的子样分布与总体真分布的差别在置信水平 90%时小于 20%. 类似地, 当要求精度好于 0.05 时, 由附表 12 可查得 n 需大于等于 600.

需要强调指出, 柯尔莫哥洛夫检验中零假设给定的分布不能包含未知参数. 当总体理论分布包含未知参数, 即为 $F_0(x, \vartheta)$ 的情形, 需要用子样 $x_1, x_2, \cdots, x_n$ 来估计参数 ϑ 得到其估计值 $\hat{\vartheta}_n$, 但利用 $F_0(x, \hat{\vartheta}_n)$ 作为理论分布得到的 D_n, D_n^+, D_n^- 统计量不再是分布自由的, 而是与 $F_0(x, \vartheta)$ 有关. 这样, 式 (12.4.32) 和式 (12.4.33) 的极限分布不再成立, 无法据此确定临界域.

例 12.11　低统计实验的拟合优度检验

作为柯尔莫哥洛夫检验的一个例子, 我们来考察一个低统计实验. 假定观测了中性 K 介子衰变为 $\pi^+ e^- \nu$ 的 30 个事例, 其衰变时间谱如图 12.9(a) 中折线 $S_{30}(t)$ 所示. 希望检验该中性 K 介子是不是 $\overline{\mathrm{K}}^0$. 当认为这些介子是 $\overline{\mathrm{K}}^0$ 时 (零假设 H_0), 衰变时间谱如图 12.9 (a) 中的连续曲线 $F_0(t)$ 所示.

从测量和理论曲线确定了

$$D_{30,\mathrm{obs}} = \max_{0<t<\infty} |S_{30}(t) - F_0(t)| = 0.17.$$

选定显著性水平 $\alpha = 0.10$, 从附表 12 查得

$$D_{30,\alpha=0.10} = 0.2176,$$

因此不能拒绝零假设 H_0.

作为对比, 利用同样的观测数据对 H_0 作皮尔逊 χ^2 检验. 将衰变时间划分为 4 个子区间：0~3, 3~5, 5~7, 7~18(单位 0.89×10^{t10}s) 如图 12.9 (b) 所示, 这样的划分满足每一子区间内观测频数 $\geqslant 5$ 的要求, 而且每一子区间内的理论概率积分值大致相等. 这时, 式 (12.4.16) 表示的统计量 X^2 渐近地服从 $\chi^2(3)$ 分布, 由 χ^2 分布表查得

$$\chi^2_{\alpha=0.10}(3) = 6.251,$$

而观测值算出

$$X^2_{\mathrm{obs}} = 3.0 < \chi^2_{0.10}(3),$$

因此, 皮尔逊检验同样得到不能拒绝原假设 H_0 的结论.

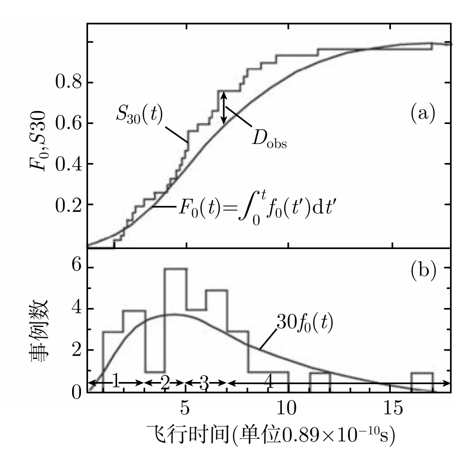


图 12.9　衰变时间谱的理论预期和实验测量值

(a) 累积分布 (柯尔莫哥洛夫检验); (b) 微分分布 (皮尔逊 χ^2 检验)

12.4.6　斯米尔诺夫–克拉美–冯 · 迈希斯检验

与柯尔莫哥洛夫检验利用子样分布函数与原假设总体累积分布函数之差的极大值作为拟合优度的检验统计量不同, 斯米尔诺夫–克拉美–冯 · 迈希斯 (Smirnov-Cramer-von Mises) 检验利用子样分布函数与原假设总体累积分布函数的偏差 W^2

作为检验统计量, 其定义为

$$W^2 = \int_{-\infty}^{\infty} \left[S_n(x) - F_0(x)\right]^2 f_0(x)\,\mathrm{d}x, \tag{12.4.38}$$

其中 $f_0(x)$ 是原假设的概率密度, $F_0(x)$ 是原假设的分布函数. 将式 (12.4.29) 的子样分布函数代入上式, 注意到 $F(-\infty) \equiv 0, F(+\infty) \equiv 1$, 可得 [35]

$$\begin{aligned} W^2 &= \int_{-\infty}^{x_1} F_0^2(x)\,\mathrm{d}F_0(x) + \sum_{i=1}^{n-1} \int_{x_i}^{x_{i+1}} \left[\frac{i}{n} - F_0(x)\right]^2 \mathrm{d}F_0(x) \\ &\quad + \int_{x_n}^{\infty} \left[1 - F_0(x)\right]^2 \mathrm{d}F_0(x) \\ &= \frac{1}{12n^2} + \frac{1}{n}\sum_{i=1}^{n} \left[F_0(x_i) - \frac{2i-1}{2n}\right]^2. \end{aligned} \tag{12.4.39}$$

该式可用于检验统计量 W^2 的具体计算.

对于一特定的 x 值, $nS_n(x)$ 为期望值 $nF_0(x)$ 的二项分布变量, 故有

$$\begin{aligned} &E\left[S_n(x)\right] = F_0(x), \\ &E\left\{\left[S_n(x) - F_0(x)\right]^2\right\} = \frac{1}{n} F_0(x)\left[1 - F_0(x)\right]. \end{aligned}$$

式 (12.4.39) 所示的统计量具有均值和方差:

$$\begin{aligned} &E\left(W^2\right) = \frac{1}{n}\int_0^1 F_0(x)\left[1 - F_0(x)\right]\mathrm{d}F_0(x) = \frac{1}{6n}, \\ &V\left(W^2\right) = E\left(W^4\right) - \left[E\left(W^2\right)\right]^2 = \frac{4n-3}{180n^3}. \end{aligned} \tag{12.4.40}$$

即使对于有限的 n 值, 统计量 W^2 的分布亦与总体 X 的分布完全无关, 这一点容易证明. 作变量代换 $Y = F_0(x)$, 统计量 W^2 变成

$$W^2 = \int_0^1 \left[S_n(Y) - Y\right]^2 \mathrm{d}Y,$$

它完全独立于原假设的概率密度 $f_0(x)$, 也即统计量 W^2 的分布与原假设的分布相独立.

斯米尔诺夫证明了, 在 $n \to \infty$ 的极限情形下, 当原假设为真, nW^2 的特征函数为

$$\varphi(t) = \left\{(2\mathrm{i}t)^{1/2} \Big/ \sin\left[(2\mathrm{i}t)^{1/2}\right]\right\}^{1/2}. \tag{12.4.41}$$

Anderson 和 Darling 通过 $\varphi(t)$ 的反演计算了 nW^2 的临界值, 对于某些常用的显著性水平 α, nW^2 临界值列于表 12.7. α 及其对应的 nW^2 临界值的更详细的表见书

末的附表 13. Marshall 证明了, nW^2 趋于其极限分布的速度很快, 以至于在 $n \geqslant 3$ 的情形下, 表 12.7 所列的 nW^2 的临界值都是精确的. 因此, 可利用此表基于统计量 W^2 进行关于原假设 $F(x) = F_0(x)$ 的检验. 由 W^2 的定义式 (12.4.38) 可知, nW^2 的临界域在其临界值的上侧.

表 12.7 斯米尔诺夫–克拉美–冯 · 迈希斯检验统计量 nW^2 的临界值

α	nW^2 的临界值
0.10	0.347
0.05	0.461
0.01	0.743
0.001	1.168

与柯尔莫哥洛夫检验相同, 斯米尔诺夫–克拉美–冯 · 迈希斯检验中零假设给定的分布不能包含未知参数, 否则统计量 nW^2 不再是分布自由的, 式 (12.4.41) 的极限分布不再成立, 无法据此确定临界域.

12.5 信号的统计显著性

如我们在第 11 章小信号测量问题中提到的, 实验测量值通常同时包含信号和本底的贡献, 而且信号和本底的测量都存在统计涨落, 即它们都是随机变量. 当我们在信号区观测到的事例数 n 明显地高于预期的本底事例数 b, 我们会判断观测到了信号事例. 例如, 在图 11.1(b) 中不变质量 $m_{\eta\pi^+\pi^-} \approx 958\text{MeV}$ 附近观测到了 η' 粒子的信号. 显然, n 比 b 大出越多, 对于 “观测到了信号事例” 这一判断的可信程度越高. 信号的**统计显著性**(statistical significance) 是物理学家对 “观测到了信号事例” 这一判断的定量化表征.

在讨论信号的统计显著性问题时, 零假设 H_0 通常表示为观察到的实验现象可以只用已知的现象或本底函数圆满地描述, 备择假设 H_1 则表示观察到的实验现象需要用已知的现象或本底函数, 加上未知、待寻找的新信号过程的贡献才能完整地描述. 信号的统计显著性就是观察到的实验现象偏离已知的现象或本底函数、发现新信号、新过程的定量表征. 信号的统计显著性越高, 发现新过程的可信度越大. 在粒子物理实验中, 基本上达成了一种共识, 如果实验中观察到一种新信号其显著性 $S \geqslant 5$, 则可以认为 “发现” 了一种新信号; 当新信号的显著性 $S \geqslant 3$ ($S \geqslant 2$), 则只能说新信号的存在有强 (弱) 的证据. 所以在寻找新现象的实验测量中, 信号的统计显著性尤为重要.

12.5.1 实验 P 值

粒子物理实验中, 对是否观测到信号事例这一判断的定量表征方法之一是给出

实验 P 值, 它定义为

$$P(u_{\text{obs}}) = P(u \geqslant u_{\text{obs}}|H_0) = \begin{cases} \displaystyle\sum_{u \geqslant u_{\text{obs}}} p_u(H_0), & u\text{离散型;} \\ \displaystyle\int_{u \geqslant u_{\text{obs}}} f(u|H_0)\mathrm{d}u, & u\text{连续型.} \end{cases} \tag{12.5.1}$$

其中, u 是实验观测量 (随机变量) 或用实验观测量构造的统计量, u_{obs} 是某个实验测量到的 u 值, $p_u(H_0)$ 是零假设 H_0 为真时离散型变量 u 的概率分布, $f(u|H_0)$ 是零假设 H_0 为真时连续型变量 u 的概率密度函数, 实验 P 值是 H_0 为真时 $u \geqslant u_{\text{obs}}$ 的概率. 不失一般性, 我们设定 u 值越大, H_0 为真的可能性越小. 那么, $P(u_{\text{obs}})$ 值就是该实验测量值 u_{obs} 与零假设 H_0 不一致性的某种定量表征, $P(u_{\text{obs}})$ 值越小, H_0 为真的可能性越小.

假定存在某种信号事例, 它的 u 值集中地出现在一个特定的区域称为信号区. 实验中信号区内观测到的信号事例数可视为期望值 s 的泊松变量, 信号区内观测到的本底事例数可视为期望值 b 的泊松变量, 则信号区内观测到总事例数 n 的概率为

$$f(n; s, b) = \frac{(s+b)^n}{n!}\mathrm{e}^{-(s+b)}. \tag{12.5.2}$$

假定实验观测到的信号区内总事例数为 n_{obs}, 令零假设 H_0 为观测到的事例仅仅是由于本底的贡献, 则相应的实验 P 值为

$$P(n_{\text{obs}}) = P(n \geqslant n_{\text{obs}}|H_0) = \sum_{n=n_{\text{obs}}}^{\infty} f(n; s=0, b) = 1 - \sum_{n=0}^{n_{\text{obs}}-1} \frac{b^n}{n!}\mathrm{e}^{-b}. \tag{12.5.3}$$

实验 P 值越小, H_0 为真 (观测到的事例仅仅是由于本底的贡献) 的可能性越小. 举一个数值例子：假定 $b = 0.6, n_{\text{obs}} = 5$, 则相应的实验 P 值约等于 4×10^{-4}, 这表示实验观测到 $n \geqslant n_{\text{obs}}$=5 个事例仅仅是由于本底涨落导致的概率小到只有 4×10^{-4}.

在实际应用中需要注意信号区内平均本底事例数 (即期望值)b 可能存在系统误差 (见第 11 章的讨论). 通常信号区内平均本底事例数是由信号区外附近的本底区 (称为边带区, sideband) 的事例数分布确定的, 这实际上隐含着两个假设：边带区内不包含信号事例, 且能相当好地反映信号区内的本底贡献. 然而实际的测量数据中边带区内观测到的事例数分布往往不够平滑 (如图 11.1), 边带区的宽度也没有严格的规则加以判定, 而是依赖于实验者对 “边带区内不包含信号事例, 且能相当好地反映信号区内的本底贡献” 的主观判断予以确定, 因此存在不确定性, 通常利用不同的本底函数形式、不同的边带区宽度来确定 b 的系统误差. 例如, 在上面的例子中, $b = 0.6$, 假定考虑了系统误差后信号区内平均本底事例数 b 的范围为 0.5~0.7, 则相应的实验 P 值范围为 $2 \times 10^{-4} \sim 8 \times 10^{-4}$. 作为最后结果, 可以报道实验 P 值的这一范围, 或者保守地仅仅报道实验 P 值为 8×10^{-4}.

另一个问题是信号区的宽度. 不同的信号区宽度给出不同的 n_{obs} 和 b 值, 从而得到不同的实验 P 值. 一般信号区的宽度取为足以包含绝大部分信号事例 (如果存在的话). 例如, 对于信号的实验分布为正态函数的情形, 信号区间取为 $x_{\text{c}} \pm 2.5\sigma$ 或 $x_{\text{c}} \pm 3\sigma$ 都是合理的选择, 这里 x_{c} 和 σ 分别是正态函数的中心值和标准偏差, 它们分别包含了全部信号事例的 98.8%和 99.7%. 报道最终结果时推荐同时给出信号区间的宽度和实验 P 值.

12.5.2 信号的统计显著性

实验 P 值表示的是实验观测到的现象仅仅是由于本底涨落导致的概率, 但是实验观测到的现象与待寻找的新现象、新信号之间的关系表达的不直观, 后者通常用信号的统计显著性 S 来表示. 然而在粒子物理实验数据分析的发展过程中, 使用了不同的统计显著性 S 的定义. 例如, 对于 12.5.1 节中讨论的信号区内观测到的信号和本底事例数都服从泊松分布的情形, 文献 [71],[72] 对于信号统计显著性 S 有如下各种定义:

$$S_1 = (n - b)/\sqrt{b}, \tag{12.5.4}$$

$$S_2 = (n - b)/\sqrt{n}, \tag{12.5.5}$$

$$S_{12} = \sqrt{n} - \sqrt{b}, \tag{12.5.6}$$

$$S_{B1} = S_1 - k(\alpha)\sqrt{n/b}, \tag{12.5.7}$$

$$S_{B12} = 2S_{12} - k(\alpha), \tag{12.5.8}$$

$$\int_{-\infty}^{S_N} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} \mathrm{d}x = \sum_{i=0}^{n-1} \frac{b^i}{i!} \mathrm{e}^{-b}, \tag{12.5.9}$$

其中 n 是信号区内观测到的总事例数, b 是信号区内观测到的本底事例数的期望值, 它被认为是一个已知值. S 的不同角标表示来源于不同的定义, $k(\alpha)$ 是一个与观察到信号事例的确定程度相关的系数, 对于它们的详细说明请阅读文献 [71],[72]. 以上 S 的不同定义都是基于实验测量中信号区内的计数来检验零假设 H_0 和备择假设 H_1, 这样的实验被称为 “计数实验”.

在许多情形下, 通过实验测量值或它的统计量的分布来检验零假设 H_0 和备择假设 H_1 比简单地通过信号区内的计数来检验要更为准确和精确, 这时, 需要用似然比方法进行检验. 假定 $L(b)$ 和 $L(s+b)$ 分别为用零假设 H_0 和备择假设 H_1 的概率密度构造的似然函数, $L_{\text{m}}(b)$ 和 $L_{\text{m}}(s+b)$ 分别为用 $L(b)$ 和 $L(s+b)$ 拟合实验数据得到的似然函数极大值, I.Narsky[72] 在 $-2\ln[L_{\text{m}}(b)/L_{\text{m}}(s+b)]$ 服从自由度为 1 的 χ^2 分布的假定下给出了信号的统计显著性

$$S = [2(\ln L_{\text{m}}(s+b) - \ln L_{\text{m}}(b))]^{1/2}. \tag{12.5.10}$$

似然函数可以取为标准的形式

$$L(s+b)=\prod_{i=1}^{N}\left[\omega_s f_s(u_i)+(1-\omega_s)f_b(u_i)\right], \tag{12.5.11}$$

$$L(b)=\prod_{i=1}^{N}f_b(u_i), \tag{12.5.12}$$

其中 N 为实验中观察到的总事例数, 需要特别指出, N 与前面提到的信号区内观测到的总事例数 n 是不同的, 因为实验观测的区域总是比信号区要宽得多, 所以 N 为总大于 n; s 和 b 分别表示实验观测的区域中的信号和本底事例数, 实验测量值或它的统计量表示为 u, u_i 为第 i 个事例的测量值, $f_s(u)$ 和 $f_b(u)$ 分别表示信号和本底的概率密度函数, ω_s 为信号事例数的权因子.

但一般取广义的似然函数更为适当, 即实验中观察到的总事例数 N 不认为是一常数而认为是一泊松变量, 这就考虑了观察事例数泊松分布的不确定性, 这时的似然函数形式为

$$L(s+b)=\frac{\mathrm{e}^{-(s+b)}}{N!}\prod_{i=1}^{N}\left[sf_s(u_i)+bf_b(u_i)\right], \tag{12.5.13}$$

$$L(b)=\frac{\mathrm{e}^{-b}b^N}{N!}\prod_{i=1}^{N}f_b(u_i). \tag{12.5.14}$$

由于实验 P 值表示的是实验数据与零假设之间的不一致性, 实验 P 值越小, 零假设为真的可能性越小, 备择假设为真的可能性越大. 因此文献 [73] 将信号显著性与实验 P 值联系起来, 对信号显著性作了如下的定义：

$$\int_{-S}^{S}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\mathrm{d}x=1-P(u\geqslant u_{\mathrm{obs}}|H_{bg})\equiv 1-P(u_{\mathrm{obs}}). \tag{12.5.15}$$

该式的左边是正态分布在 $\pm S$ 个标准差 $(\pm S\sigma)$ 内的积分概率. 在这样的定义下, 相对应的 S 值和实验 P 值见表 12.8.

对于计数实验的情形, 实验 P 值为

$$P(n_{\mathrm{obs}})=P(n\geqslant n_{\mathrm{obs}}|H_0)=\sum_{n=n_{\mathrm{obs}}}^{\infty}\frac{b^n}{n!}\mathrm{e}^{-b}=1-\sum_{n=0}^{n_{\mathrm{obs}}-1}\frac{b^n}{n!}\mathrm{e}^{-b}. \tag{12.5.16}$$

代入式 (12.5.15), 立即有

$$\int_{-S}^{S}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\mathrm{d}x=\sum_{n=0}^{n_{\mathrm{obs}}-1}\frac{b^n}{n!}\mathrm{e}^{-b}. \tag{12.5.17}$$

与式 (12.5.9) 比较, 积分下限有所不同.

表 12.8 实验 P 值与统计显著性 S 的对应关系

S 值	实验 P 值
1	0.3173
2	0.0455
3	0.0027
4	6.3×10^{-5}
5	5.7×10^{-7}
6	2.0×10^{-9}

对于似然比方法的情形, 如前所述, 假定零假设 H_0 是实验数据仅仅由本底似然函数 $L(b)$ 描述, 备择假设 H_1 是实验数据由 $L(s+b)$ 描述, 假定 $L(b)$ 有 m 个待定参数 (描述本底概率密度 $f_b(u)$ 的参数), $L(s+b)$ 有 $k(>m)$ 个待定参数 (描述本底概率密度 $f_b(u)$ 和信号概率密度 $f_s(u)$ 的参数以及它们之间的相对权重因子). 按照 12.2.2 节的叙述, H_0(似然函数 $L(b)$) 是对 H_1(似然函数 $L(s+b)$) 中 k 个待定参数中的 r 个参数加以固定, 或加上了 r 个约束条件; 或者说, r 是备择假设 H_1 与零假设 H_0 的待估计的独立参数个数之差, 或自由度之差, 于是似然比统计量

$$\lambda = L_{\rm m}(b)/L_{\rm m}(s+b) \tag{12.5.18}$$

是零假设 H_0 的合理的检验统计量, 这里 $L_{\rm m}(b)$ 是利用 $L(b)$ 拟合实验数据得到的极大似然函数值. 文献 [35] 证明了, 当 H_0 为真, 在子样容量很大的情形下, 统计量

$$u = -2\ln\lambda = 2(\ln L_{\rm m}(s+b) - \ln L_{\rm m}(b)) \tag{12.5.19}$$

渐近地服从 $\chi^2(r)$ 分布. 当 λ 的观测值接近于 1, H_0 为真的可能性很大; 当 λ 的观测值接近于 0, H_0 为真的可能性很小. 所以, λ 的临界域在 λ 值接近 0 的区域, 相应地, u 的临界域在 u 值大的区域. 倘若 u 的实验观测值为 $u_{\rm obs}$, 则由式 (12.5.1) 知实验 P 值为

$$P(u_{\rm obs}) = \int_{u_{\rm obs}}^{\infty} \chi^2(u;r){\rm d}u, \tag{12.5.20}$$

代入式 (12.5.15) 立即得到利用似然比统计量计算信号统计显著性 S 的表达式

$$\int_{-S}^{S} \frac{1}{\sqrt{2\pi}}\,{\rm e}^{-x^2/2}{\rm d}x = 1 - P(u_{\rm obs}) = \int_0^{u_{\rm obs}} \chi^2(u;r){\rm d}u. \tag{12.5.21}$$

对于 $r=1$ 的特殊情形, 有

$$\int_{-S}^{S} \frac{1}{\sqrt{2\pi}}\,{\rm e}^{-x^2/2}{\rm d}x = \int_0^{u_{\rm obs}} \chi^2(u;1){\rm d}u = 2\int_0^{\sqrt{u_{\rm obs}}} \frac{1}{\sqrt{2\pi}}{\rm e}^{-x^2/2}{\rm d}x,$$

立即可得

$$S = \sqrt{u_{\rm obs}} = [2(\ln L_{\rm m}(s+b) - \ln L_{\rm m}(b))]^{1/2}, \tag{12.5.22}$$

它与 I. Narsky 的式 (12.5.10) 完全一致.

由以上讨论可知, 式 (12.5.15) 定义的信号统计显著性无论对于计数实验或似然比方法都是适用的, 因此避免了多重定义. 利用似然比统计量计算信号统计显著性的式 (12.5.21) 的优点在于避免了确定信号区和本底边带区时带来的不确定因素, 但由以上的推导过程可知必须满足大子样容量的要求. 对于小子样容量的情形, 仍然应当用式 (12.5.21) 计算信号统计显著性 S.

由于 S 值有多种定义, 因此实验结果的报道中, 在陈述某个信号的统计显著性 S 值为多大时, 应当说明其明确的定义.

12.6 独立性检验

第 3 章中我们提到, n 维随机向量与一维随机变量的最大区别, 在于前者的性质不仅与各分量有关, 而且依赖于各分量间的相互关联. 当各分量间相互独立, 随机向量实际上是互不关联的 n 个一维随机变量, 问题的处理就大大简化. 独立性检验就是利用随机向量的子样观测值来决定各分量是否相互独立的问题. 独立性检验最常见的是 χ^2 检验法.

12.6.1 二维随机变量分量的独立性检验

设二维随机变量 $\{X,Y\}$ 的分布函数及边沿分布函数分别是 $F(x,y)$ 及 $F_X(x)$, $F_Y(y)$. 由 3.1 节所述, 随机变量 X,Y 相互独立的条件是对一切 x 和 y, 下式成立：

$$F(x,y)=F_X(x)\cdot F_Y(y).$$

因此, 两个随机变量之间独立性检验的零假设和备择假设可表示为

$$\begin{aligned}H_0: F(x,y)=F_X(x)\cdot F_Y(y),\\ H_1: F(x,y)\neq F_X(x)\cdot F_Y(y).\end{aligned}$$

设有一个二维总体, 随机地抽取容量 n 的二维子样 $(x_1,\ \ y_1)$, $\cdots$, $(x_n,\ \ y_n)$. 将 X 和 Y 的值域分别划分为 I 和 J 个子区间, 子样中 X 属于第 i 子区间, Y 值属于第 j 子区间的事例数记为 n_{ij}, 并记

$$n_{i\cdot}\equiv\sum_{j=1}^{J}n_{ij},\quad n_{\cdot j}\equiv\sum_{i=1}^{I}n_{ij}. \tag{12.6.1}$$

子样容量 n 值满足归一化条件

$$\sum_{i=1}^{I}\sum_{j=1}^{J}n_{ij}=\sum_{i=1}^{I}n_{i\cdot}=\sum_{j=1}^{J}n_{\cdot j}=n. \tag{12.6.2}$$

这些数目可方便地写成**联列表**的形式如表 12.9 所示.

表 12.9 二维随机变量的联列表

n_{ij} j / i	1	2	3	$\cdots$	J	$n_{i\cdot}$
1	n_{11}	n_{12}	n_{13}	$\cdots$	n_{1J}	$n_{1\cdot}$
2	n_{21}	n_{22}	n_{23}		n_{2J}	$n_{2\cdot}$
3	n_{31}	n_{32}	n_{33}		n_{3J}	$n_{3\cdot}$
$\vdots$	$\vdots$					$\vdots$
I	n_{I1}	n_{I2}	n_{I3}		n_{IJ}	$n_{I\cdot}$
$n_{\cdot j}$	$n_{\cdot 1}$	$n_{\cdot 2}$	$n_{\cdot 3}$	$\cdots$	$n_{\cdot J}$	n

从总体$\{X, Y\}$中任意抽取一个元素, 它的 x 值落入第 i 子区间, y 值落入第 j 个子区间这一事件的概率记为 p_{ij}, 以 $p_{i.}$ 和 $p_{j.}$ 分别表示相应的边沿概率, 则有

$$p_{i\cdot} = \sum_{j=1}^{J} p_{ij}, \quad p_{\cdot j} = \sum_{i=1}^{I} p_{ij}, \tag{12.6.3}$$

$$\sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij} = \sum_{i=1}^{I} p_{i\cdot} = \sum_{j=1}^{I} p_{\cdot j} = 1. \tag{12.6.4}$$

这样, 独立性检验的零假设可表示为

$$H_0 : p_{ij} = p_{i\cdot}p_{\cdot j}, \quad i = 1, \cdots, I; \quad j = 1, \cdots, J. \tag{12.6.5}$$

由于条件式 (12.6.4) 的存在, I 个 $p_{i\cdot}$ 和 J 个 $p_{\cdot j}$ 中只有 $(I-1)$ 个 $p_{i\cdot}$ 和 $(J-1)$ 个 $p_{\cdot j}$ 是独立的. $p_{i\cdot}$ 和 $p_{\cdot j}$ 可用极大似然法估计. 当 H_0 为真, 似然函数为

$$\begin{aligned} L &= \prod_{i=1}^{I}\prod_{j=1}^{J}(p_{ij})^{n_{ij}} = \prod_{i=1}^{I}\prod_{j=1}^{J}(p_{i\cdot})^{n_{ij}}(p_{\cdot j})^{n_{ij}} = \prod_{i=1}^{I}(p_{i\cdot})^{n_{i\cdot}}\prod_{j=1}^{J}(p_{\cdot j})^{n_{\cdot j}} \\ &= \left(1 - \sum_{i=1}^{I-1} p_{i\cdot}\right)^{n_{I\cdot}} \left(1 - \sum_{j=1}^{J-1} p_{\cdot j}\right)^{n_{\cdot J}} \left(\prod_{i=1}^{I-1} p_{i\cdot}^{n_{i\cdot}}\right)\left(\prod_{j=1}^{J-1} p_{i\cdot}^{n_{\cdot j}}\right). \end{aligned}$$

由此得

$$\ln L = n_{I\cdot}\ln\left(1 - \sum_{i=1}^{I-1} p_{i\cdot}\right) + n_{\cdot J}\ln\left(1 - \sum_{j=1}^{J-1} p_{\cdot j}\right) + \sum_{i=1}^{I-1} n_{i\cdot}\ln p_{i\cdot} + \sum_{j=1}^{J-1} n_{\cdot j}\ln p_{\cdot j}.$$

由似然方程

$$\begin{cases} \dfrac{\partial \ln L}{\partial p_{i\cdot}} = 0, \quad i = 1, \cdots, I-1, \\ \\ \dfrac{\partial \ln L}{\partial p_{\cdot j}} = 0, \quad j = 1, \cdots, J-1. \end{cases}$$

求出

$$\begin{cases} p_{i\cdot} = n_{i\cdot}/A, \quad i = 1, \cdots, I, \\ p_{\cdot j} = n_{\cdot j}/B, \quad j = 1, \cdots, J. \end{cases}$$

其中,

$$A \equiv n_{I\cdot}/p_{I\cdot}, \ B \equiv n_{\cdot J}/p_{\cdot J}.$$

从式 (12.6.4) 得到

$$\sum_{i=1}^{I} p_{i\cdot} = \frac{\sum\limits_{i=1}^{I} n_{i\cdot}}{A} = \frac{n}{A} = 1,$$

$$\sum_{j=1}^{I} p_{\cdot j} = \frac{\sum\limits_{j=1}^{J} n_{\cdot j}}{B} = \frac{n}{B} = 1,$$

故有 $A = B = n$. 最后解出 $p_i.$ 和 $p_{\cdot j}$ 的极大似然估计值为

$$\begin{cases} \hat{p}_{i\cdot} = n_{i\cdot}/n, \quad i = 1, \cdots, I-1 \\ \hat{p}_{\cdot j} = n_{\cdot j}/n, \quad j = 1, \cdots, J-1 \end{cases} \tag{12.6.6}$$

因此, H_0 为真时, 据式 (12.6.5), 应有

$$\hat{p}_{ij} = \hat{p}_{i\cdot} \cdot \hat{p}_{\cdot j} = n_i.n_{\cdot j}/n^2.$$

现在 $n\hat{p}_{ij}$ 可作为 “理论频数”, n_{ij} 可作为观测频数, 当 H_0 为真, 这两者的差别将很小. 类似于 12.4.2 节皮尔逊 χ^2 检验, 我们可构作检验统计量 $X^{\,2}$

$$X^2 = \sum_{i=1}^{I}\sum_{j=1}^{J} \frac{(n_{ij} - n_i.n_{\cdot j}/n)^2}{n_i.n_{\cdot j}/n} = n\left(\sum_{i=1}^{I}\sum_{j=1}^{J} \frac{n_{ij}^2}{n_i.n_{\cdot j}} - 1\right). \tag{12.6.7}$$

当 H_0 为真, 并且对所有的 i, j, n_{ij} 充分大, 则 X^2 渐近地服从 χ^2 分布, 自由度 ν 等于独立观测数和独立的未知参数个数之差. 在 $I \cdot J$ 个观测数 n_{ij} 中, 由于归一化条件 $\sum\limits_{i,j} n_{ij} = 1$ 的存在, 只有 $IJ-1$ 个是独立的; 未知参数 $\hat{p}_{i\cdot}$ 和 $\hat{p}_{j\cdot}$ 中只有 $I-1$ 个和 $J-1$ 个是独立的 (见式 (12.6.6)), 因此, 自由度 ν 等于

$$\nu = (IJ-1) - [(I-1) + (J-1)] = (I-1)(J-1) = IJ - (I+J) + 1. \tag{12.6.8}$$

这样, 给定显著性水平 α 后, 可从 $\chi^2(\nu)$ 分布表查得临界值 $\chi^2_\alpha(\nu)$. 当观测值 $X^2_{\text{obs}} > \chi^2_\alpha(\nu)$, 则在显著性水平 α 上拒绝零假设 H_0.

例 12.12 动量分量的独立性

图 12.10 是一个二维散点图, 它表示 19GeV/c 的质子–质子碰撞中产生的 670 个 Λ 超子, 在质子–质子的质心系中, 纵向动量 p_l 和横向动量 p_t 的分布. 所谓纵向动量 p_l, 是指 Λ 超子动量在入射质子方向上的动量分量; p_t 则为垂直于该方向的动量分量, 要求根据这些信息来检验 p_l 与 p_t 是否相互独立.

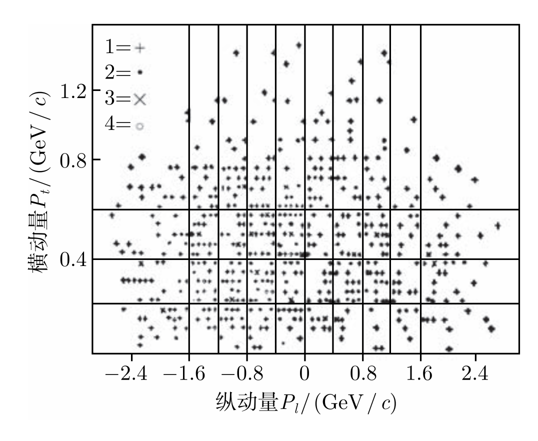


图 12.10 Λ 超子的纵向动量 p_l 和横向动量 p_t 的分布

图中数字表示该动量值处 Λ 超子的计数. + 表示数字 1

将横动量 p_t 覆盖的范围划分为 4 个互不相容的子区间 (标以 A_{1-4}); 纵动量 p_l 覆盖的范围划分为 10 个子区间 (标以 B_{1-10}), 如图 12.10 中相互垂直的线所划分的区域所示. 得出联列表, 见表 12.10.

表 12.10 Λ 超子纵向动量和横向动量数据的联列表

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9	B_{10}	$n_{i\cdot}$
A_1	20	8	13	12	13	11	7	5	7	10	106
A_2	20	23	22	26	15	14	21	20	9	21	191
A_3	21	22	15	24	35	10	19	22	8	15	191
A_4	20	18	25	19	25	25	19	18	6	7	182
$n_{\cdot j}$	81	71	75	81	88	60	66	65	30	53	670

按表中的数据和式 (12.6.7) 算得 $X^2_{\text{obs}} = 39.8$. 由式 (12.6.8) 可知, 自由度为 $(4-1)\times(10-1) = 27$. 当给定显著性水平 $\alpha = 0.05$, 临界值 $\chi^2_{\alpha=0.05}(27) = 40.11$. 故在显著性水平 $\alpha = 0.05$ 上不能拒绝零假设, 即纵动量 p_l 与横动量 p_t 可认为相互独立.

12.6.2　多维随机变量分量的独立性检验

多维随机变量分量的独立性问题比二维随机变量分量的独立性问题要复杂得多, 本节就三维随机变量的情形进行讨论[87], 更高维随机变量分量的独立性问题可以参照这些方法进行检验.

设有一个三维总体 $\{X,Y,Z\}$, 最重要的独立性为：

(1) X,Y,Z 三者相互独立;

(2) 其中的两个变量联合地独立于第三个变量;

(3) 给定一个变量值的条件下, 另两个变量的条件独立性.

其中独立性 (1) 最强, (2) 次之, (3) 最弱.

将三维总体 $\{X,Y,Z\}$ 的 X,Y 和 Z 的值域分别划分为 I,J 和 K 个子区间, 从总体$\{X, Y, Z\}$中任意抽取一个元素, 它的 x 值落入第 i 子区间, y 值落入第 j 个子区间, z 值落入第 k 个子区间这一事件的概率记为 p_{ijk}, 并记相应的边沿概率为

$$p_{ij\cdot} \equiv \sum_{k=1}^{K} p_{ijk}, \quad p_{\cdot jk} \equiv \sum_{i=1}^{I} p_{ijk}, \quad p_{i\cdot k} \equiv \sum_{j=1}^{J} p_{ijk}. \tag{12.6.9}$$

$$p_{i\cdot\cdot} \equiv \sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}, \quad p_{\cdot j\cdot} \equiv \sum_{i=1}^{I}\sum_{k=1}^{K} p_{ijk}, \quad p_{\cdot\cdot k} \equiv \sum_{i=1}^{I}\sum_{j=1}^{J} p_{ijk}. \tag{12.6.10}$$

根据归一化条件有

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk} = \sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij\cdot} = \sum_{i=1}^{I}\sum_{k=1}^{K} p_{i\cdot k} = \sum_{j=1}^{J}\sum_{k=1}^{K} p_{\cdot jk} = \sum_{i=1}^{I} p_{i\cdot\cdot} = \sum_{j=1}^{J} p_{\cdot j\cdot} = \sum_{k=1}^{K} p_{\cdot\cdot k} = 1. \tag{12.6.11}$$

可以证明, 对应于上述三类独立性, 有表 12.11 所示的概率表达式.

表 12.11　三维随机变量的各种独立性及其概率表达式

独立性	记号	独立性成立的概率表达式
X,Y,Z 相互独立	(X,Y,Z)	$p_{ijk} = p_{i\cdot\cdot}p_{\cdot j\cdot}p_{\cdot\cdot k} \quad \forall i,j,k$
X,Y 联合独立于 Z	$(XY,\ Z)$	$p_{ijk} = p_{ij\cdot}p_{\cdot\cdot k} \quad \forall i,j,k$
Y,Z 联合独立于 X	$(YZ,\ X)$	$p_{ijk} = p_{\cdot jk}p_{i\cdot\cdot} \quad \forall i,j,k$
Z,X 联合独立于 Y	$(ZX,\ Y)$	$p_{ijk} = p_{i\cdot k}p_{\cdot j\cdot} \quad \forall i,j,k$
给定 X, Y,Z 条件独立	$(XY,\ XZ)$	$p_{ijk} = p_{ij\cdot}p_{i\cdot k}/p_{i\cdot\cdot} \quad \forall i,j,k$
给定 Y, Z,X 条件独立	$(YZ,\ XY)$	$p_{ijk} = p_{ij\cdot}p_{\cdot jk}/p_{\cdot j\cdot} \quad \forall i,j,k$
给定 Z, X,Y 条件独立	$(ZX,\ YZ)$	$p_{ijk} = p_{i\cdot k}p_{\cdot jk}/p_{\cdot\cdot k} \quad \forall i,j,k$

为了检验各独立性是否成立, 就需要检验概率 p_{ijk} 是否满足相应的概率表达式.

设有一个三维总体, 随机地抽取容量 n 的三维子样 (x_1, y_1, z_1), $\cdots$, (x_n, y_n, z_n). 将 X,Y 和 Z 的值域分别划分为 I,J 和 K 个子区间, 子样中 X 属于第 i 子区间, Y 值属于第 j 子区间, Z 值属于第 k 子区间的事例数记为 n_{ijk}, 并记

$$n_{ij\cdot} \equiv \sum_{k=1}^{K} n_{ijk}, \quad n_{\cdot jk} \equiv \sum_{i=1}^{I} n_{ijk}, \quad n_{i\cdot k} \equiv \sum_{j=1}^{J} n_{ijk}. \tag{12.6.12}$$

$$n_{i\cdot\cdot} \equiv \sum_{j=1}^{J}\sum_{k=1}^{K} n_{ijk}, \quad n_{\cdot j\cdot} \equiv \sum_{i=1}^{I}\sum_{k=1}^{K} n_{ijk}, \quad n_{\cdot\cdot k} \equiv \sum_{i=1}^{I}\sum_{j=1}^{J} n_{ijk}. \tag{12.6.13}$$

子样容量 n 值满足归一化条件

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} n_{ijk} = \sum_{i=1}^{I} n_{i\cdot\cdot} = \sum_{j=1}^{J} n_{\cdot j\cdot} = \sum_{k=1}^{K} n_{\cdot\cdot k} = n. \tag{12.6.14}$$

在独立性为真的条件下, 对于一切可能的 i, j, k 值, 事例数 n_{ijk} 的理论频数应为

$$m_{ijk} = np_{ijk}. \tag{12.6.15}$$

将独立性成立的各概率表达式右边的各个概率用相应的频率去估计, 便得到相应的理论频数估计 $\hat{m}_{ijk} = n\hat{p}_{ijk}$. 类似于 12.4.2 节皮尔逊 χ^2 检验, 我们可构建检验统计量 X^2

$$X^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \frac{(n_{ijk} - \hat{m}_{ijk})^2}{\hat{m}_{ijk}}. \tag{12.6.16}$$

当独立性假设 H_0 为真, 并且对所有的 i, j, k, n_{ijk} 充分大, 则 X^2 渐近地服从 χ^2 分布, 自由度 ν 等于独立观测数和独立的未知参数个数之差. 当独立性假设为真, X^2 值应当偏小, 即临界域在上侧.

对于 (X,Y,Z) 独立性检验问题, $p_{i\cdot\cdot}$, $p_{\cdot j\cdot}$ 和 $p_{\cdot\cdot k}$ 的估计值由下式求得：

$$\hat{p}_{i\cdot\cdot} = \frac{n_{i\cdot\cdot}}{n}, \quad \hat{p}_{\cdot j\cdot} = \frac{n_{\cdot j\cdot}}{n}, \quad \hat{p}_{\cdot\cdot k} = \frac{n_{\cdot\cdot k}}{n}. \tag{12.6.17}$$

由于条件式 (12.6.14) 的存在, 独立观测数为 $(IJK - 1)$ 个. 根据式 (12.6.11), 未知参数 I 个 $p_{i\cdot\cdot}$, J 个 $p_{\cdot j\cdot}$ 和 K 个 $p_{\cdot\cdot k}$ 中只有 $(I-1)$ 个 $p_{i\cdot\cdot}$, $(J-1)$ 个 $p_{\cdot j\cdot}$ 和 $(K-1)$ 个 $p_{\cdot\cdot k}$ 是独立的, 因此, 自由度 ν 等于

$$\nu = (IJK - 1) - [(I-1) + (J-1) + (K-1)] = IJK - (I + J + K) + 2. \tag{12.6.18}$$

对于 (XY, Z) 独立性检验问题, $p_{ij\cdot}$ 和 $p_{\cdot\cdot k}$ 的估计值由下式求得

$$\hat{p}_{ij\cdot} = \frac{n_{ij\cdot}}{n}, \quad \hat{p}_{\cdot\cdot k} = \frac{n_{\cdot\cdot k}}{n}. \tag{12.6.19}$$

根据式 (12.6.11), 未知参数 IJ 个 $p_{ij\cdot}$ 和 K 个 $p_{\cdot\cdot k}$ 中只有 $(IJ-1)$ 个 $p_{ij\cdot}$ 和 $(K-1)$ 个 $p_{\cdot\cdot k}$ 是独立的, 因此, 自由度 ν 等于

$$\nu = (IJK-1) - [(IJ-1)+(K-1)] = (IJ-1)+(K-1). \qquad (12.6.20)$$

这样, 给定显著性水平 α 后, 可从 $\chi^2(\nu)$ 分布表查得临界值 $\chi^2_\alpha(\nu)$. 当观测值 $X^2_{\text{obs}} > \chi^2_\alpha(\nu)$, 则在显著性水平 α 上拒绝零假设 H_0.

表 12.12 汇总了三维随机变量独立性检验所需的各种量的表示式.

表 12.12 三维随机变量独立性检验所需的各种量

独立性	p_{ijk}概率表达式	理论频数估计 $\hat{m}_{ijk}$	X^2渐近分布自由度ν
(X,Y,Z)	$p_{i\cdot\cdot}p_{\cdot j\cdot}p_{\cdot\cdot k}$	$n_{i\cdot\cdot}n_{\cdot j\cdot}n_{\cdot\cdot k}/n^2$	$IJK-(I+J+K)+2$
$(XY,\ Z)$	$p_{ij\cdot}p_{\cdot\cdot k}$	$n_{ij\cdot}n_{\cdot\cdot k}/n$	$(IJ-1)+(K-1)$
$(YZ,\ X)$	$p_{\cdot jk}p_{i\cdot\cdot}$	$n_{\cdot jk}n_{i\cdot\cdot}/n$	$(JK-1)+(I-1)$
$(ZX,\ Y)$	$p_{i\cdot k}p_{\cdot j\cdot}$	$n_{i\cdot k}n_{\cdot j\cdot}/n$	$(IK-1)+(J-1)$
$(XY,\ XZ)$	$p_{ij\cdot}p_{i\cdot k}/p_{i\cdot\cdot}$	$n_{ij\cdot}n_{i\cdot k}/n_{i\cdot\cdot}$	$I(J-1)(K-1)$
$(YZ,\ XY)$	$p_{ij\cdot}p_{\cdot jk}/p_{\cdot j\cdot}$	$n_{ij\cdot}n_{\cdot jk}/n_{\cdot j\cdot}$	$J(I-1)(K-1)$
$(ZX,\ YZ)$	$p_{i\cdot k}p_{\cdot jk}/p_{\cdot\cdot k}$	$n_{i\cdot k}n_{\cdot jk}/n_{\cdot\cdot k}$	$K(I-1)(J-1)$

由于三维随机变量 3 类独立性之间的强弱关系, 对其独立性的检验, 在给定的显著性水平下, 可先检验原假设 $H_0:(X,Y,Z)$. 若接受 H_0, 则其他两类独立性均成立. 否则, 进一步检验原假设 $(XY,\ Z)$, $(YZ,\ X)$ 或 $(ZX,\ Y)$, 若其中两种为真则第三类独立性均成立; 否则继续检验第三类独立性.

对于更高维随机变量各分量的独立性问题, 可用类似的方法进行检验.

12.7 相关性检验

与独立性检验利用随机向量的子样观测值来决定各分量是否相互独立不同, 相关性检验要解决的问题是随机向量各分量是否存在相互关联, 如果相互关联, 其关联程度有多大.

12.7.1 Pearson 相关系数的检验

在 3.3 节中我们已经提到, 随机变量 X 和 Y 的 Pearson 相关系数 ρ 定义为

$$\rho \equiv \frac{\text{cov}\,(X,Y)}{\sigma_X\sigma_Y},$$

它是随机变量 X 和 Y 关联程度的定量表述. 相关系数的检验问题一般为①两个随机变量间是否存在相关性, 即 ρ 是否等于 0; ②如果存在相关性, 其相关系数是否等于某个特定常数.

这是一个参数检验问题, 关键是要构建一个检验统计量, 它的分布为已知, 与待检验的参数有简单的关系, 它又是随机变量 X 和 Y 的观测值 (样本值) 的函数, 通常称为枢轴量. 这样可利用该枢轴量来进行统计检验.

(1) Pearson 相关系数 $\rho = 0$ 的检验

设有二维总体 $\{X, Y\}$, 随机地抽取容量 n 的二维子样 $(X_1, Y_1), \cdots, (X_n, Y_n)$. 子样相关系数为 (见式 (6.2.12))

$$r = \frac{S_{XY}}{S_X S_Y} = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\left[\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 \sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2\right]^{1/2}},$$

其中 $\bar{X}, \bar{Y}$ 是随机变量 X, Y 各自的子样平均. 可以证明, 当 n 充分大, 子样相关系数 r 是随机变量 X 和 Y 的相关系数 ρ 的好的估计量. 同时, 可以证明, 当 n 充分大, 统计量

$$t = r\sqrt{\frac{n-2}{1-r^2}} \tag{12.7.1}$$

服从自由度 $n-2$ 的 t 分布. 因此 t 可以作为相关系数 ρ 的检验统计量. 给定置信水平 $\gamma \equiv 1-\alpha$, 可以进行原假设 $H_0 : \rho = 0$ 的双侧检验或单侧检验:

(a) 双侧检验

原假设 $H_0 : \rho = 0$

备择假设 $H_1 : \rho \neq 0$

原假设 H_0 的临界域: $|t_{\text{obs}}| > t_{\alpha/2}(n-2)$

其中, 统计量 t 的观测值由样本观测值按式 (12.7.1) 计算求得:

$$t_{\text{obs}} = r_{\text{obs}}\sqrt{\frac{n-2}{1-r_{\text{obs}}^2}}$$

而双侧 α 分位数 $t_{\alpha/2}(n-2)$ 由 t 分布表查得. 于是可由观测值 t_{obs} 决定原假设是否成立.

(b) 单侧检验

$$H_0 : \rho \geqslant 0, \quad H_1 : \rho < 0,$$

原假设 H_0 的临界域为: $t_{\text{obs}} < -t_{\alpha}(n-2)$;

$$H_0 : \rho \leqslant 0, H_1 : \rho > 0,$$

原假设 H_0 的临界域为: $t_{\text{obs}} > t_{\alpha}(n-2)$.

查表得单侧 α 分位数 $t_\alpha(n-2)$, 由观测值 t_{obs} 决定原假设是否成立.

(2) Pearson 相关系数 $\rho=\rho_0$ 的检验

当两个随机变量间存在相关性, 即 ρ 不等于 0, 那么一个自然的问题就是其相关性的强弱及相关系数的符号 (即正相关还是负相关). 这时待检验的假设应是 ρ 是否等于某个特定的常数 ρ_0, 即检验的原假设是 $H_0:\rho=\rho_0$.

同样, 问题的关键在于找到一个适当的检验统计量. Fisher 提出了基于 r 变换成 z_r 后构建的检验统计量

$$Z=\frac{z_r-z_{\rho_0}}{1/\sqrt{n-3}} \tag{12.7.2}$$

其中,

$$z_r=\frac{1}{2}\ln\frac{1+r}{1-r},\quad z_{\rho_0}=\frac{1}{2}\ln\frac{1+\rho_0}{1-\rho_0}. \tag{12.7.3}$$

这一变换称为 Fisher 变换[22]. 这里, z_r 是 r 的单调增函数, 如图 12.11 所示 (类似地, z_{ρ_0} 是 ρ_0 的单调增函数). 可以证明, 当 n 充分大时, 统计量 Z 渐近于 $N(0,1)$ 分布[23]. 于是式 (12.7.2) 的 Z 就是原假设 $H_0:\rho=\rho_0$ 的一个适当的检验统计量.

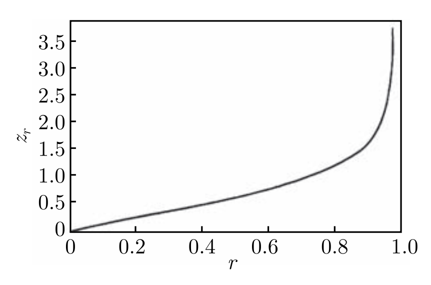


图 12.11 z_r 与 r 的函数关系

给定置信水平 $\gamma\equiv1-\alpha$, 对于双侧检验, 我们有

原假设 H_0: $H_0:\rho=\rho_0$,

备择假设 H_1: $H_0:\rho\neq\rho_0$,

原假设 H_0 的临界域: $|Z_{\text{obs}}|>\lambda_{\alpha/2}$.

其中 $\lambda_{\alpha/2}$ 为标准正态的双侧 α 分位数. 这样, 由观测值 $Z_{\text{obs}}=\dfrac{z_r-z_{\rho_0}}{1/\sqrt{n-3}}$ 可决定原假设是否成立.

对于单侧检验, 有

H_0: $\rho\geqslant\rho_0$ H_0: $\rho<\rho_0$,

原假设 H_0 的临界域为：$Z_{\text{obs}}<-\lambda_\alpha$.

H_0: $\rho\leqslant\rho_0$ H_1: $\rho>\rho_0$,

原假设 H_0 的临界域为：$Z_{\text{obs}}>\lambda_\alpha$.

其中 λ_α 为标准正态的上侧 α 分位数. 于是由观测值 Z_{obs} 可决定原假设是否成立. 对于下侧检验, 可用类似的方法进行检验.

值得注意的是, 当样本量 $n<25$ 时, 用 Fisher 提出的式 (12.7.2) 的统计量 Z 作为服从 $N(0,1)$ 分布的随机变量容易发生较大偏差, 因此, 它通常应用于 $n \geqslant 25$ 的场合. 对于样本量 $n \geqslant 10$ 的情形, 统计学家建议利用如下的检验统计量

$$Z^* = \left(z_r^* - z_{\rho_0}^*\right)\sqrt{n-1}. \tag{12.7.4}$$

该统计量近似于 $N(0,1)$ 分布, 其中,

$$z_r^* = z_r - \frac{3z_r + r}{4n}, \quad z_{\rho_0}^* = z_{\rho_0} - \frac{3z_{\rho_0} + \rho_0}{4n}, \tag{12.7.5}$$

这里, z_r^* 同样是 r 的单调增函数.

给定置信水平 $\gamma \equiv 1-\alpha$, 对双侧检验, 有

原假设　　$H_0:\ \ \rho = \rho_0,$

备择假设　　$H_1:\ \ \rho \neq \rho_0,$

原假设 H_0 的临界域　　$|Z_{\text{obs}}^*| > \lambda_{\alpha/2}.$

由观测值 $Z_{\text{obs}}^* = \left(z_r^* - z_{\rho_0}^*\right)\sqrt{n-1}$ 决定原假设是否成立.

对单侧检验, 由于统计量 Z^* 中的 z_r^* 同样是 r 的单调增函数, 故可用类似于上述的方法作单侧检验:

对备择假设 $H_1:\ \ \rho<0$, 原假设 H_0 的临界域为 $Z_{\text{obs}}^* < -\lambda_\alpha$;

对备择假设 $H_1:\ \ \rho>0$, 原假设 H_0 的临界域为 $Z_{\text{obs}}^* > \lambda_\alpha$.

12.7.2　Spearman 秩相关检验

当随机变量 X 和 Y 的精确观测值不可得而只知道它们的大小次序时, 变量 X 和 Y 的相关性这时不可能用 Pearson 相关系数 ρ 来表示, 而需要用 Spearman 秩相关系数 r_S 来表示 [24] .

(1) Spearman 秩相关检验的基本方法

设有二维总体 $\{X,Y\}$, 随机地抽取容量 n 的二维子样 $(X_1,\ Y_1),\ \cdots,\ (X_n,\ Y_n)$. 令 R_i 是 X_i 在 $(X_1,\cdots,X_n)$ 中的秩 (顺序号), Q_i 是 Y_i 在 $(Y_1,\cdots,Y_n)$ 中的秩. 如果 X_i 与 Y_i 具有同步性, 则 R_i 与 Q_i 也应表现出同步性; 反之亦然.

仿照样本相关系数的计算方法,

$$r = \frac{S_{XY}}{S_X S_Y} = \frac{\sum\limits_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\left[\sum\limits_{i=1}^{n}(X_i-\overline{X})^2\sum\limits_{i=1}^{n}(Y_i-\overline{Y})^2\right]^{1/2}},$$

可定义 Spearman 秩相关系数：

$$r_S = \frac{\sum_{i=1}^{n}(R_i - \overline{R})(Q_i - \overline{Q})}{\left[\sum_{i=1}^{n}(R_i - \overline{R})^2 \sum_{i=1}^{n}(Q_i - \overline{Q})^2\right]^{1/2}}. \tag{12.7.6}$$

这里平均秩为

$$\bar{R} = \frac{1}{n}\sum_{i=1}^{n} R_i, \bar{Q} = \frac{1}{n}\sum_{i=1}^{n} Q_i. \tag{12.7.7}$$

注意到

$$\sum_{i=1}^{n} R_i = \sum_{i=1}^{n} Q_i = \frac{n\,(n+1)}{2},$$

$$\sum_{i=1}^{n} R_i^2 = \sum_{i=1}^{n} Q_i^2 = \frac{n\,(n+1)\,(2n+1)}{6},$$

故 r_S 可简化为

$$r_S = 1 - \frac{6}{n\,(n^2-1)}\sum_{i=1}^{n}(R_i - Q_i)^2. \tag{12.7.8}$$

r_S 的范围为

$$-1 \leqslant r_S \leqslant +1.$$

r_S 的绝对值指示相关性的强度, 其符号则表示相关性的方向, 例如正号对应于秩 R_i 增大时 Q_i 也增大; 负号对应于秩 R_i 增大时 Q_i 减小.

应当特别注意 Pearson 相关系数 ρ 与 Spearman 秩相关系数 r_S 含义的差别, 前者描述的是随机变量 X 与 Y 之间的相关性, 后者描述的是随机变量 X 和 Y 样本的顺序统计量 R 与 Q 之间的相关性, 即 X 和 Y 的秩相关性. 不知道样本值而只知道样本值顺序的情形下, 只能作 Spearman 秩相关检验; 知道样本值的情形下, Pearson 相关性和 Spearman 秩相关性两种检验都可以进行. 但是对于同一组测量样本, 相关系数 ρ 与 Spearman 秩相关系数 r_S 会是不相同的. 一个明显的例子是对于随机变量 X 和 $Y = X^2$, $\rho_{XY} = 0$ 而 $r_S = 1$.

为了进行假设检验, 需要知道 r_S 的分布. 可以证明, 对于 n >10 的情形, 统计量

$$t_S = r_S\sqrt{\frac{n-2}{1-r_S^2}} \tag{12.7.9}$$

在 H_0 为真的条件下 (随机变量 X 和 Y 秩不相关, 即 $r_S = 0$) 近似于 $t\,(n-2)$ 分布[25], 这样, 就可利用 t_S 进行相关性检验.

对于双侧检验

$$H_0: r_S = 0, \quad H_1: r_S \neq 0,$$

给定显著性水平 α, 原假设 H_0 的临界域为

$$|t_{S,\text{obs}}| > t_{\alpha/2}(n-2).$$

由观测值 $t_{S,\text{obs}} = r_{S,\text{obs}}\sqrt{\dfrac{n-2}{1-r_{S,\text{obs}}^2}}$ 决定原假设 (X 和 Y 秩不相关) 是否成立.

对于单侧检验

$$H_0: r_S \leqslant 0, \quad H_1: r_S > 0,$$

原假设 H_0(X 和 Y 秩负相关) 的临界域为

$$t_{S,\text{obs}} > t_\alpha(n-2).$$

用类似的方法可进行 $H_0: r_S \geqslant 0$, $H_1: r_S < 0$ 的单侧检验, 即原假设 H_0(X 和 Y 秩正相关) 的临界域为

$$t_{S,\text{obs}} < -t_\alpha(n-2).$$

对于小样本 (例如 n <10) 的情形, 统计量 t_S 的分布不能再用 $t(n-2)$ 分布作为近似. 这种情形下, 原假设 H_0 的双侧和单侧检验的临界域临界值 $t_{\alpha/2}$ 和 t_α 可查书末所附的附表 14[25].

(2) 大样本近似

H_0 为真条件下 ($r_S = 0$), r_S 的分布对于 0 对称, 且有如下性质:

$$E_{H_0}(r_S) = 0, \quad V_{H_0}(r_S) = \frac{1}{n-1}. \tag{12.7.10}$$

可以证明, 大样本近似下, 有

$$z_S = r_S\sqrt{n-1} \sim N(0,1) \tag{12.7.11}$$

故当 $n \geqslant 25$(有些书籍认为至少要满足 $n \geqslant 100$)[25], 可利用 z_S 进行秩相关性的检验. 于是, 原假设 H_0(随机变量 X 和 Y 秩不相关) 的双侧检验临界域为

$$|z_{S,\text{obs}}| > \lambda_{\alpha/2}.$$

而单侧检验临界域为

$$z_{S,\text{obs}} < -\lambda_\alpha \quad (H_0: r_S \geqslant 0, \quad H_1: r_S < 0);$$

$$z_{S,\text{obs}} > \lambda_\alpha \quad (H_0: r_S \leqslant 0, \quad H_1: r_S > 0).$$

其中 $\lambda_{\alpha/2}$ 和 λ_α 分别为标准正态分布的双侧和上侧 α 分位数.

利用式 (12.7.11) 进行秩相关性检验的好处在于, 对于 $r_S = \pm 1$ 的情形, 式 (12.7.9) 的统计量 t_S 没有意义而无法计算, 而统计量 z_S 是可以计算的, 因而可用来进行假设检验.

(3) "结" 的处理

如果数据中有重复数据, 例如, $(X_1, \cdots, X_n)$ 中 l 个 X_i 值相等, 它们相应的秩值 $R_{m+1}, \cdots, R_{m+l}$ 相等, 这时, 每个秩值都等于这 l 个秩值的平均秩: $m+(l+1)/2$. 这 l 个秩形成一个 "结".

令 $\tau_i(X), \tau_j(Y)$ 为 $(X_1, \cdots, X_n)$ 和 $(Y_1, \cdots, Y_n)$ 中的结统计量, 即 $\tau_i(X)$ 为 $(X_1, \cdots, X_n)$ 中第 i 个结的长度 (结中的样本数). 这时秩相关系数 r_S 需修改为 $\tilde{r}_S$

$$\tilde{r}_S = \frac{\dfrac{n(n^2-1)}{6} - \dfrac{1}{12}\left[\sum_i \left(\tau_i^3(X) - \tau_i(X)\right) + \sum_j \left(\tau_j^3(Y) - \tau_j(Y)\right)\right] - \sum_{i=1}^{n}(R_i - Q_i)^2}{2\sqrt{\left[\dfrac{n(n^2-1)}{12} - \dfrac{1}{12}\sum_i \left(\tau_i^3(X) - \tau_i(X)\right)\right]\left[\dfrac{n(n^2-1)}{12} - \dfrac{1}{12}\sum_j \left(\tau_j^3(Y) - \tau_j(Y)\right)\right]}}. \tag{12.7.12}$$

这时, 统计量

$$\tilde{t}_S = \tilde{r}_S\sqrt{\frac{n-2}{1-\tilde{r}_S^2}} \tag{12.7.13}$$

在 H_0 为真条件下 (随机变量 X 和 Y 不相关) 遵从 $t(n-2)$ 分布, 于是, 就可利用 $\tilde{t}_S$ 进行秩相关性的检验.

当 "结" 的长度较小时, $\tilde{r}_S$ 的分布仍可用无结时的 r_S 分布作为近似.

当样本容量 n 大时, 可用极限分布

$$\tilde{z}_S = \tilde{r}_S\sqrt{n-1} \sim N(0,1) \tag{12.7.14}$$

进行大样本检验.

12.7.3 Kendall τ 相关检验

Kendall 于 1938 年提出一种与 Spearman 秩相关检验相似的检验法[26]. 基本思想是从样本对 $(X_i, Y_i), i = 1, 2, \cdots, n$ 是否协同一致的角度来检验变量 (X,Y) 之间的相关性.

所谓的协同性 (concordance) 是指, 当数对 (X_i, Y_i) 与 (X_j, Y_j) 满足关系式

$$(X_j - X_i)(Y_j - Y_i) > 0, \quad \forall j > i, i, j = 1, \cdots, n, \tag{12.7.15}$$

称 (X_i, Y_i) 与 (X_j, Y_j) 具有协同性, 这表示 (X_i, Y_i) 与 (X_j, Y_j) 变化方向相同; 当满足关系式

$$(X_j - X_i)(Y_j - Y_i) < 0, \quad \forall j > i, i, j = 1, \cdots, n. \tag{12.7.16}$$

则数对 (X_i, Y_i) 与 (X_j, Y_j) 不协同, 或 (X_i, Y_i) 与 (X_j, Y_j) 变化方向相反. 故协同性表示的是前后两个数对的秩大小的变化是同向还是反向.

全部样本对 $(X_i, Y_i), i = 1, 2, \cdots, n$ 可能有的前后数对共有

$$N_t = \begin{pmatrix} n \\ 2 \end{pmatrix} = \frac{n(n-1)}{2}. \tag{12.7.17}$$

令 N_C 为同向数对的数目, N_D 为反向数对的数目, 且

$$S \equiv N_C - N_D, \tag{12.7.18}$$

Kendall 相关系数定义为

$$\tau = \frac{S}{N_t} = \frac{2S}{n(n-1)}. \tag{12.7.19}$$

注意到 $N_t = N_C + N_D$, 上式亦可表示为

$$\tau = 1 - \frac{4N_D}{n(n-1)}, \tag{12.7.20}$$

$$\tau = \frac{4N_C}{n(n-1)} - 1. \tag{12.7.21}$$

在 $N_C = N_t$(即全部数对同向) 的情形下, 显然 $\tau = 1$; 与此相反, 当 $N_D = N_t$(全部数对反向), $\tau = -1$, 因此有

$$\tau \in [-1, 1].$$

$\tau > 0$ 对应于 (X, Y) 秩正相关, $\tau < 0$ 对应于 (X,Y) 秩负相关, 而 $\tau = 0$ 即 $N_C = N_D$(同向、反向数对相等), 则 (X,Y) 秩不相关.

为了算得 Kendall 相关系数 τ 的具体数值, 必须知道 N_C 或 N_D 的值. 假定 X_i 已从小到大排序, 即 $X_i < X_{i+1}, i = 1, \cdots, n-1$. 令 $Y_1, \cdots, Y_n$ 的秩为 $R_1, \cdots, R_n$, 记

$$p_i = \sum_{j>i} I(R_j > R_i), \quad i = 1, \cdots, n. \tag{12.7.22}$$

$$q_i = \sum_{j>i} I(R_j < R_i), \quad i = 1, \cdots, n. \tag{12.7.23}$$

其中 $I(R_j > R_i)$ 表示 $(R_j > R_i)$ 时 $I(R_j > R_i) = 1$; $I(R_j < R_i)$ 表示 $(R_j < R_i)$ 时 $I(R_j < R_i) = 1$, 则有

$$N_C = \sum_{i=1}^{n} p_i, \quad N_D = \sum_{i=1}^{n} q_i. \tag{12.7.24}$$

代入式 (12.7.19)~ 式 (12.7.21) 即可求得 τ 值.

可以证明, 在原假设 H_0 成立, 即 (X,Y) 不相关的条件下, 统计量 τ 的分布有以下性质[27]：

(1) τ 的期望值和方差为

$$E_{H_0}(\tau) = 0, \quad V_{H_0}(\tau) = \frac{2(2n+5)}{9n(n-1)}. \tag{12.7.25}$$

(2) τ 的分布关于原点 0 对称.

(3) $n \to \infty$ 时

$$\tau/\sigma_{H_0}(\tau) \sim N(0,1). \tag{12.7.26}$$

实际上, 当 n >10, 正态分布就是合理的近似.

如果数据中有重复数据, 例如 $(X_1, \cdots, X_n)$ 中 τ_X 个 X_i 值相等, 它们相应的秩值相等, 这时, 每个秩值都等于这 τ_X 个秩值的平均秩, 这 τ_X 个秩形成一个 “结”. 这时 Kendall 相关系数修正为

$$\tilde{\tau} = \frac{S}{\sqrt{n(n-1)/2 - T_X}\sqrt{n(n-1)/2 - T_Y}}. \tag{12.7.27}$$

式中,

$$T_X = \frac{1}{2}\sum^{N_X}\left(\tau_X^2 - \tau_X\right), \quad T_Y = \frac{1}{2}\sum^{N_Y}\left(\tau_Y^2 - \tau_Y\right), \tag{12.7.28}$$

这里 N_X, N_Y 分别为 $\{X_i\}$ 中结的个数和 $\{Y_i\}$ 中结的个数, τ_X, τ_Y 分别为 $\{X_i\}$ 的结的长度 (结中的样本数) 和 $\{Y_i\}$ 的结长.

根据式 (12.7.26) 统计量 τ 的分布, Kendall 相关检验方法如下. 大样本情形下 (n >10), 给定置信水平 $\gamma = 1 - \alpha$, 对于双侧检验, 原假设 $H_0(\tau = 0)$ 的临界域为：

$$|\tau_{\text{obs}}/\sigma_{H_0}(\tau)| > \lambda_{\alpha/2}.$$

对于单侧检验, 原假设 H_0 的临界域为

$$\tau_{\text{obs}}/\sigma_{H_0}(\tau) < -\lambda_\alpha (H_0 : r_S \geqslant 0, H_1 : r_S < 0);$$

$$\tau_{\text{obs}}/\sigma_{H_0}(\tau) > \lambda_\alpha (H_0 : r_S \leqslant 0, H_1 : r_S > 0).$$

其中 $\lambda_{\alpha/2}$ 和 λ_{α} 分别为标准正态分布的双侧和上侧 α 分位数.

对于小样本 (n <10) 的情形, 统计量 τ 的分布不能再用正态分布作为近似. 这种情形下, 原假设 H_0 的双侧和单侧检验的临界域临界值 $\tau_{\alpha/2}$ 和 τ_{α} 可查书末所附的附表 15[25].

关于 Kendall 相关系数 τ与 Spearman 秩相关系数 r_S 间的关系, 显然, 对于同样一组数据样本对 $(X_i, Y_i), i = 1, 2, \cdots, n$, Kendall 秩相关检验和 Spearman 秩相关检验都可以进行, 但实际中 Spearman 秩相关检验应用较多, 主要原因是它的计算比较简单. 另一个原因是, 当数据样本对 $(X_i, Y_i), i = 1, 2, \cdots, n$ 来自于二维正态分布, r_S值是 Pearson 相关系数 ρ_{XY} 的好的近似, 而 τ 则不具备这一性质. 但统计量 τ 也有自身的优点, 当样本容量增大时, τ 的样本分布逼近正态分布比 r_S 要快得多 [28], 因而, 应用正态近似进行 Kendall秩相关检验时样本容量可以比较小, 而 Spearman 秩相关检验的样本容量则必须充分大.

虽然 τ 与 r_S 都是同一组数据样本对的秩相关性的表示, 但两者的尺度不尽相同. 一般而言, 不可能由 τ 推导出 r_S 值, 反之亦然. 但是 τ 的绝对值总是小于 r_S 的绝对值 (除非 τ 和 r_S 的绝对值都等于 1); 当样本容量增大, 比值 $\tau/r_S \to 0.67$. 尽管如此, 这两个统计量本质上应用了关于数据的同样多的信息, 因此对于同一组数据样本对的秩相关性是否为 0, 两种检验方法会得出相同的结论.

例 12.13 体重与肺活量的相关性

10 名男生的体重 x 和肺活量 y 的数据如下:

x	75	95	85	70	76	68	60	66	80	88
y	2.62	2.91	2.94	2.11	2.17	1.98	2.04	2.20	2.65	2.78

给定显著性水平 $\alpha = 0.10$, 利用 Pearson 相关系数、Spearman 相关系数和 Kendall 相关系数检验肺活量与体重是否相关.

解

(1) Pearson 相关系数检验

问题的原假设和备择假设为

$$H_0: \ \rho = 0, H_1: \ \rho \neq 0,$$

由体重 x 和肺活量 y 的数据, 按式 (6.2.12)计算得到 Pearson 子样相关系数的观测值为 $r_{\text{obs}} = 0.8887$, 根据式 (12.7.1) 算得 $t_{\text{obs}} = r_{\text{obs}}\sqrt{\dfrac{n-2}{1-r_{\text{obs}}^2}} = 5.483$. 查 t 分布的双侧 α 分位数表 8得 $t_{0.05}(8) = 1.860$, 故 t_{obs} 落入原假设的临界域内, 应拒绝原假设, 即肺活量与体重有相关性.

(2) Spearman 相关系数检验

问题的原假设和备择假设为

$$H_0 : r_S = 0, \quad H_1 : r_S \neq 0,$$

将 x 从小到大排序后有

$R_x = i$	1	2	3	4	5	6	7	8	9	10
$R_y = R_i$	2	5	1	3	6	4	7	10	8	9

由式 (12.7.8) 计算 Spearman 相关系数得

$$r_{S,\text{obs}} = 1 - \frac{6}{n\left(n^2-1\right)}\sum_{i=1}^{n}\left(R_x - R_y\right)^2 = 0.8424.$$

再由式 (12.7.9) 计算 $t_{S,\text{obs}} = r_{S,\text{obs}}\sqrt{\dfrac{n-2}{1-r_{S,\text{obs}}^2}} = 4.422$, 大于 t 分布的双侧 α 分位数 $t_{0.05}(8) = 1.860$; 或由表 13 查得 Spearman 相关系数临界值为 0.564, $r_{S,\text{obs}} = 0.8424 > 0.564$, 故拒绝原假设, 即肺活量与体重有相关性.

(3) Kandall 相关系数检验

问题的原假设和备择假设为

$$H_0 : \tau = 0, \quad H_1 : \tau \neq 0$$

根据 (2) 中 $R_y = R_i$ 的序列, 可得

$I(R_j > R_i)$	8	5	7	6	4	4	3	0	1	0
$I(R_j < R_i)$	1	3	0	0	1	0	0	2	0	0

由式 (12.7.22)~ 式 (12.7.24) 知 $N_C = 38$, $N_D = 7$, 由式 (12.7.19) 算得

$$\tau_{\text{obs}} = \frac{2S}{n(n-1)} = \frac{2(38-7)}{90} = 0.689.$$

据式 (12.7.25) 有

$$\sigma_{H_0}(\tau) = \sqrt{\frac{2(2n+5)}{9n(n-1)}} = 0.2485, \quad \left|\tau_{\text{obs}}/\sigma_{H_0}(\tau)\right| = 2.77,$$

查标准正态分布累积分布函数表 6 得 $\lambda_{\alpha/2} = 1.645$. 故 $\tau_{\text{obs}}/\sigma_{H_0}(\tau)$ 值落在原假设 H_0 的临界域, 即拒绝原假设 H_0 而接受备择假设肺活量和体重相关. 或由表 14 查得 Kandall 秩相关检验临界值为 0.511, 而 $\tau_{\text{obs}} = 0.689 > 0.511$, 故结论相同.

由此可见, 对于这一特定问题, 三种检验方法得出相同的结论.

12.7.4 多变量 Kendall 协和系数检验

前两小节介绍的 Spearman 和 Kendall 秩检验都针对两个变量间的相关性. 这种相关的概念可延拓至多变量间的相关性. 例如为了诊断病情, 通常要对病人做多项检查, 这些结果之间是否存在相关性? 为了确定粒子的种类, 可以测定粒子的飞行时间 TOF, 粒子在介质中的电离能量损失 $\mathrm{d}E/\mathrm{d}x$, 粒子在量能器中的沉积能量 E_{dep} 等, 它们的数值之间是否存在相关性?

Kendall 和 Babington[29] 于 1939 年提出的多变量协和系数 (coefficient of concordance) 检验就是针对这类多变量间的相关性问题的.

设有 m 维随机向量 X 的 n 个样本, 每个样本有 m 个变量, 即 $\underset{\sim}{X}=\{X_{ij}\}_{n\times m}$, 其中 i 标记样本号, j 标记变量 (分量) 号. 并定义变量 j 的向量为 $\boldsymbol{X}_j=(X_{1j}, X_{2j},\cdots,X_{nj})$. 记 r_{ij} 为 X_{ij} 在 $X_{1j},X_{2j},\cdots,X_{nj}$ 中的秩, 表示成如下的数据表形式:

表 12.13 多变量的秩表示

	变量 1	变量 2	$\cdots$	变量 m	秩和
秩	r_{11}	r_{12}		r_{1m}	$R_{1\cdot}$
	r_{21}	r_{22}		r_{2m}	$R_{2\cdot}$
	$\vdots$	$\vdots$		$\vdots$	$\vdots$
	r_{n1}	r_{n2}		r_{nm}	$R_{n\cdot}$

记

$$R_{i\cdot}=\sum_{j=1}^{m}r_{ij},\quad j=1,\cdots,m, \tag{12.7.29}$$

其中 $R_{i\cdot}$ 是第 i 个样本的秩和值. 注意到

$$R_{\cdot j}\equiv\sum_{i=1}^{n}r_{ij}=1+\cdots+n=\frac{n\,(n+1)}{2},\quad j=1,\cdots,m, \tag{12.7.30}$$

并且 $R_{\cdot j}$ 与 j 无关, 于是总秩和值

$$R=\sum_{i=1}^{n}\sum_{j=1}^{m}r_{ij}=\sum_{i=1}^{n}R_{i\cdot}=\sum_{j=1}^{m}R_{\cdot j}=\frac{mn\,(n+1)}{2}. \tag{12.7.31}$$

设变量 $\boldsymbol{X}_j$ 与 $\boldsymbol{X}_k$ 之间有相关性, 则 $X_{1j},X_{2j},\cdots,X_{nj}$ 的值的大小顺序, 即秩 $r_{1j},r_{2j},\cdots,r_{nj}$ 与 $X_{1k},X_{2k},\cdots,X_{nk}$ 的秩 $r_{1k},r_{2k},\cdots,r_{nk}$ 有一致性, 例如若 $r_{1j}>r_{2j}$, 即有 $r_{1k}>r_{2k}$, 等等. 反之, 当变量 $\boldsymbol{X}_j$ 与 $\boldsymbol{X}_k$ 不相关, 两者的秩就应该是随机的. 对于 m 个变量的情形, 当 m 个变量不相关 (设定为原假设 H_0), 各样本秩和值 $R_{1\cdot},R_{2\cdot},\cdots,R_{n\cdot}$ 应相近; 在 H_1 成立条件下 (m 个变量有相关性), 各个变

量的秩应有一致性, 因而某样本秩和值可能较大, 而另一样本秩和值可能较小, 即各秩和值 $R_{1\cdot}, R_{2\cdot}, \cdots, R_{n\cdot}$ 可能相差较大.

Kendall 协和系数 W 定义为样本秩和值 $R_{i\cdot}$ 的方差 S 与样本秩和值 $R_{i\cdot}$ 的最大可能方差 T 之比值 $W = S/T$, 其中,

$$S = \sum_{i=1}^{n} R_{i\cdot}^2 - \frac{R^2}{n} = \sum_{i=1}^{n} R_{i\cdot}^2 - \frac{m^2 n\,(n+1)^2}{4}.$$

$$T = \frac{m^2 n\left(n^2-1\right)}{12}.$$

故 Kendall 协和系数 W 为

$$W = \frac{S}{T} = \frac{12\sum_{i=1}^{n} R_{i\cdot}^2 - 3m^2 n\,(n+1)^2}{m^2 n\,(n^2-1)}. \tag{12.7.32}$$

W 的可能值在 [0,1] 之间. $W=0$ 表示 n 个数据样本之间各变量的秩不一致, 这意味着各变量之间不相关; $W=1$ 则表示 n 个数据样本之间各变量的秩完全一致, 意味着各变量之间完全相关. 需要注意的是, W 不可能为负值.

为了对各变量间是否存在相关性进行假设检验, 需要知道 W 的分布. Kendall 指出, 当 $n \to \infty$ 时有近似关系

$$m\,(n-1)\,W \to \chi^2(n-1). \tag{12.7.33}$$

这样, 对于较大的 n, 可用极限分布进行检验.

现在, 假设检验问题为

H_0 : n 个样本的 m 个变量秩不相关;

H_1 : n 个样本的 m 个变量秩相关.

用式 (12.7.32) 计算 W. 给定显著性水平 α, 查表得 $\chi_\alpha^2(n-1)$ 值. 当 $m\,(n-1)\,W > \chi_\alpha^2(n-1)$, 则落入 H_0 临界阈, 即接收 H_1 ; 否则, 接收 H_0. 对于较小的 n, 极限分布 $\chi^2(n-1)$ 不能描述 W 的分布, 这种情形下, 原假设 H_0 的临界域临界值可查书末所附的附表 16[25].

当样本中存在结时, 需要用平均秩方法确定秩, 这时 W 需修正为 W_c

$$W_c = \frac{12\sum_{i=1}^{n} R_i^2 - 3m^2 n\,(n+1)^2}{m^2 n\,(n^2-1) - m\sum_{j=1}^{m} T_j}, \tag{12.7.34}$$

其中

$$T_j = \sum_{l=1}^{g_j} \left(\tau_l^3 - \tau_l\right), \tag{12.7.35}$$

g_j 是 $\boldsymbol{X}_j = (X_{1j}, X_{2j}, \cdots, X_{nj})$, $j = 1, \cdots, m$ 中结的总数, τ_l 是第 l 个结的结长.

需要特别指出, Kendall 协和系数 W 的检验只应用于 $m \geqslant 3$ 的情形, 而不能用于 $m = 2$(两个变量) 的情形. 描述两个变量间相关性的 Kendall 相关系数 τ 和 Spearman 秩相关系数 r_S 均能描述负的相关性, 而 W 则不能. 事实上, 当 r_S 的值为 1, −1, 0 的时候, $m = 2$ 的 W 值分别为 1, 0, 0.5.

例 12.14 鸟的翼长、体长、嘴长间的相关性

现测得某种鸟的翼长 ($\boldsymbol{X}_1$)、体长 ($\boldsymbol{X}_2$)、嘴长 ($\boldsymbol{X}_3$) 如表 12.14 所示, 检验这 3 者是否相关.

表 12.14 某种鸟的翼长、体长和嘴长数据

	翼长 ($\boldsymbol{X}_1$/cm)		体长 ($\boldsymbol{X}_2$/cm)		嘴长 ($\boldsymbol{X}_3$/cm)		$R_{i\cdot}$
鸟编号	数据	秩	数据	秩	数据	秩	秩和
1	41	7.5	55.7	8	8.6	7.5	23
2	43	9	56.3	9	9.2	9	27
3	39.5	4	54.5	4	8	5.5	13.5
4	38	1	54.2	1.5	5.6	1	3.5
5	40.5	6	55.1	6	6.8	2	14
6	41	7.5	55.4	7	8	5.5	20
7	40	5	54.5	4	8.6	7.5	16.5
8	38.5	2	54.2	1.5	7.4	3.5	7
9	44	10	56.9	10	9.8	10	30
10	39	3	54.5	4	7.4	3.5	10.5
R							165

解

$$\sum_{j=1}^{m} T_j = \sum_{l=1}^{g_j} \left(\tau_l^3 - \tau_l\right) = \left(2^3 - 2\right) + \left(2^3 - 2\right) + \left(3^3 - 3\right) + 3 \cdot \left(2^3 - 2\right) = 54.$$

$$W_c = \frac{12 \sum_{i=1}^{n} R_i^2 - 3m^2 n \left(n+1\right)^2}{m^2 n \left(n^2 - 1\right) - m \sum_{j=1}^{m} T_j} = \frac{12 \times 657.5}{8910 - 3 \times 54} = 0.9019.$$

给定检验水平 $\alpha = 0.05$

$$m\left(n-1\right) W_c = 3 \times 9 \times 0.9019 = 24.35 > 16.92 = \chi^2_{0.05}(n-1).$$

故接受 H_1, 即某种鸟的翼长、体长、嘴长相关.

12.8 一致性检验

实验工作者经常遇到一致性检验的问题. 例如, 我们利用某种加速器的粒子流研究粒子反应, 为了得到粒子反应的充分多的事例, 数据的收集要持续相当长的时间. 要想获得可靠的实验结论, 必须保证实验条件在收集数据的过程中是始终相同的, 这可以通过收集到的数据的一致性检验得到反映. 如果某些实验条件不是始终一致, 观测值就会产生系统的偏移, 应当对这种偏移作适当的修正. 又比如不同的实验对同一个物理量作了测定, 如果要从各测定数据求出该物理量的合并测量值, 则首先要对各实验的子样测定值作一致性检验, 以保证它们之间不存在系统的差别.

下面介绍的一致性检验方法都是分布自由的, 即所使用的检验统计量与总体分布无关, 因而对任何总体都适用.

对于单个总体 X 根据其一组子样观测值来确定它的 p 分位数 (定义见 2.4 节) 是否等于指定常数这样的问题, **符号检验**提供了一种简单易行的方法. 总体的均值可以视为 “特殊” 的 p 分位数, 因而也适用此法. 这是一种 “位置” 检验问题, 只对总体 p 分位数的位置感兴趣, 而不关心总体分布函数的形式. 符号检验还可用来比较两个总体 p 分位数的一致性; **威尔科克森符号秩和检验**[20] 可应用于同样的目的, 而且后者更充分地利用了测量数据提供的信息, 因而检验更灵敏也更可靠.

12.4 节拟合优度检验解决的问题是单个总体的分布是否与原假设规定的函数形式一致, 本节所要讨论的是多个总体分布的一致性, 即它们是否具有相同的分布函数. 于是原假设可表示为

$$H_0 : f_1(x) = f_2(x) = \cdots .$$

这里 $f(x)$ 是待检验总体的概率密度. 检验的重点是各总体分布函数是否相同, 至于 $f(x)$ 的具体形式怎样我们并不关心.

对于两个子样的一致性比较, 我们介绍三种不同的检验方法. 最简单的是**游程检验**, 如果怀疑两个子样的一致性, 可首先用游程检验来决定是否排除零假设 H_0. 如果游程检验不能作出结论, 则应利用**斯米尔诺夫检验**和**威尔科克森秩和检验**[20], 它们的计算虽较游程检验复杂, 但对数据信息的利用比较充分, 因而检验两个子样之间的不一致性的效能更强. 对于多于两个子样之间的一致性检验问题, 当总体为连续函数, 可用**克鲁斯卡尔–瓦列斯**[18] 检验法; 当总体为离散型, 则应用 χ^2 检验.

在某种条件下, 利用同一组观测值, 游程检验与皮尔逊 χ^2 检验相互独立 (一般基于同一组数据作不同的检验是互相不能独立的), 这一不寻常的性质使得游程检

验成为拟合优度的皮尔逊 χ^2 检验的有益补充, 这一点在 12.4 节中已经提到.

12.8.1 符号检验

(1) 单子样问题的符号检验

单个子样的分布拟合问题在皮尔逊 χ^2 检验和柯尔莫哥洛夫检验中已经做了介绍. 但在有些场合, 我们不关心总体分布是否为一特定的函数形式, 而是需要了解总体分布的分位数 (特别是中位数) 或均值 (可视为 "特殊" 的分位数) 是否为某一指定值. 分位数和中位数的定义见式 (2.4.7)$\sim$ 式 (2.4.11). 对于这种 "位置" 检验问题, 符号检验非常适用.

设 $F(x)$ 为连续总体 X 的累积分布函数, X_p 是总体 X 的 p 分位数, 即满足

$$F(X_p) = p.$$

$X_1, \cdots, X_n$ 是 X 的随机子样, 要根据随机子样的观测值来确定, 总体 X 的 p 分位数是否等于指定常数 X_p^0. 因此, 待检验的零假设为

$$H_0 : X_p = X_p^0, \quad 0 < X_p^0 < 1. \tag{12.8.1}$$

用 z_i 表示 $X_i - X_p^0$ 的符号, 即

$$z_i \equiv \text{sign}(X_i - X_p^0) = \begin{cases} 1 & \text{若} X_i - X_p^0 \geqslant 0, \\ -1, & \text{若} X_i - X_p^0 < 0, \end{cases} \quad i = 1, 2, \cdots, n,$$

则有

$$P\{z_i = 1\} = P\left\{X_i \geqslant X_p^0\right\} = 1 - F(X_p^0) \equiv 1 - p_0,$$
$$P\{z_i = -1\} = P\left\{X_i < X_p^0\right\} = F(X_p^0) \equiv p_0,$$

其中 $p_0 \equiv F(X_p^0)$. 当 H_0 为真, 应有

$$P\{z_i = 1\} = 1 - p, \quad P\{z_i = -1\} = p.$$

令 r 为子样 $X_1, X_2, \cdots, X_n$ 中小于 X_p^0(即 $z_i = -1$) 的个数, 显然 r 是服从二项分布 $B(r; n, p_0)$ 的随机变量; 当 H_0 为真, 应有 $r \sim B(r; n, p_{)}$. 这样, 检验零假设 $X_p = X_p^0$ 的问题化简为检验假设 $p = p_0$ 的问题, 或者 r 的期望值 $E(r) = np_0$ 等于 np 的问题. 当 H_0 为真, $p = p_0$, r 的观测值远离 np 的可能性很小; 反之, H_0 不为真, r 值远离 np 的可能性变大. 于是 r 可作为零假设 H_0 的检验统计量.

临界域的确定有赖于备择假设 H_1 的选择. H_1 可以是 $X_p < X_p^0$, 或 $X_p > X_p^0$, 或 $X_p \neq X_p^0$. 对于 $X_p < X_p^0$ 的情形, 有

$$p = F(X_p) < F(X_p^0) = p_0,$$

即 $np < np_0$, 二项分布随机变量 r 的观测值倾向于比 np 大. 反之, 当 $X_p > X_p^0$, 则有

$$p = F(X_p) > F(X_p^0) = p_0,$$

r 的观测值倾向于比 np 小. 因此, 根据参数假设检验的一般方法 (见 12.2 节) 和二项分布的性质, 给定显著性水平 α, 对于不同的备择假设, 临界域可按下法求得:

(a) $H_0 : X_p = X_p^0, \quad H_1 : X_p > X_p^0$. 临界域为 $r \leqslant r_\alpha$, r_α 满足

$$\sum_{r=0}^{r_\alpha} B(r;\ n,\ p_0) \leqslant \alpha < \sum_{r=0}^{r_\alpha+1} B(r;\ n,\ p_0). \tag{12.8.2}$$

(b) $H_0 : X_p = X_p^0, H_1 : X_p < X_p^0$. 临界域为 $r \geqslant r_\alpha$, r_α 满足

$$\sum_{r=r_\alpha}^{n} B(r;\ n,\ p_0) \leqslant \alpha < \sum_{r=r_\alpha-1}^{n} B(r;\ n,\ p_0)$$

或

$$\sum_{r=0}^{r_\alpha-1} B(r;\ n,\ p_0) \geqslant 1-\alpha > \sum_{r=0}^{r_\alpha-2} B(r;\ n,\ p_0). \tag{12.8.3}$$

(c) $H_0 : X_p = X_p^0, H_1 : X_p \neq X_p^0$. 临界域为 $r \leqslant r_1$ 和 $r \geqslant r_2$, r_1, r_2 满足

$$\sum_{r=0}^{r_1} B(r;\ n,\ p_0) \leqslant \frac{\alpha}{2} < \sum_{r=0}^{r_1+1} B(r;\ n,\ p_0),$$

$$\sum_{r=r_2}^{n} B(r;\ n,\ p_0) \leqslant \frac{\alpha}{2} < \sum_{r=r_2-1}^{n} B(r;\ n,\ p_0), \tag{12.8.4}$$

或

$$\sum_{r=0}^{r_2-1} B(r;\ n,\ p_0) \geqslant 1-\frac{\alpha}{2} > \sum_{r=0}^{r_2-2} B(r;\ n,\ p_0).$$

当检验总体分布的中位数 $X_{1/2}$ 是否等于某一特定值 $X_{1/2}^0$ 时, 当 H_0 为真, 应有

$$P\{z_i = -1\} = p = \frac{1}{2} = P\{z_i = 1\} = 1-p.$$

这时, 式 (12.8.2)∼ 式 (12.8.4) 中的 p 代之以 1/2 即可.

对于子样的一组实际观测值 $x_1, x_2, \cdots, x_n$, 其中小于 X_p^0 的观测值个数记为 r_{obs}, r_{obs} 落在水平 α 的临界域内, 则在水平 α 上拒绝零假设 H_0; 反之则没有理由否定零假设, 即可认为总体的 p 分位数与给定值 X_p^0 一致.

由第 4 章的讨论知道, 当子样容量 n 很大时, 二项分布的计算很繁复, 可以应用近似性质. 当 p 接近于 0 或 1, 二项分布可用参数 $\mu = np$ 的泊松分布作为近似;

p 不接近于 0 或 1, 而 n 很大的情形下, 均值 $\mu = np$, 方差 $np(1-p)$ 的正态分布是二项分布的很好近似. 这时, 可利用标准正态变量

$$z = \frac{r - np_0}{\sqrt{np_0(1-p_0)}} \tag{12.8.5}$$

来确定临界域, 我们有

$H_1 : X_p > X_p^0$, 临界域 $z < -z_\alpha$;

$H_1 : X_p < X_p^0$, 临界域 $z > z_\alpha$;

$H_1 : X_p \neq X_p^0$, 临界域 $z < -z_{\alpha/2}$ 和 $z > z_{\alpha/2}$.

其中 z_α 为标准正态分布的上侧 α 分位数. 当

$$z_{\text{obs}} = \frac{r_{\text{obs}} - np_0}{\sqrt{np_0(1-p_0)}}$$

落在临界域内, 则在水平 α 上拒绝零假设 H_0.

作为符号检验应用的一个例子, 我们来研究束流强度随时间变化的问题. 用加速器粒子流来研究某种粒子反应, 为了在一定的时间间隔内获取较多的事例, 希望粒子流强高. 但另一方面, 实验装置可承受的粒子流强有一定的限制. 因此, 要求实验期间加速器束流强度维持在某个最佳值 μ_0 附近. 如果在实验过程中流强对于 μ_0 有明显的偏离, 则应调整加速器, 以对流强做相应的调整.

选定显著性水平 α=0.2. 实验期间在 20 个相等的时间间隔内对粒子反应的次数作计数 (它对应于加速器流强), 其中小于最佳反应次数的个数为 r_{obs}. 对于容量 20 的子样作中位数的双侧检验, 根据式 (12.8.4), 从附表 2 的累积二项分布查得

$$r_1 = 6, \quad r_2 = 14.$$

如果 $r_{\text{obs}} \leqslant 6$ 或 $r_{\text{obs}} \geqslant 14$, 则束流流强显著地偏离最佳流强, 应当做适当调整.

(2) 两子样问题的符号检验

有些场合下, 我们感兴趣的问题是比较两个总体的 p 分位数是否一致, 也可以利用符号检验, 方法与单子样问题的符号检验有许多类似之处.

设两个总体 X 和 Y 分别有连续的累积分布函数 $F(x)$ 和 $G(y)$. 从两个总体中成对地抽取容量 n 的子样, 得到 n 对数据 $(X_1, Y_1), \cdots, (X_n, Y_n)$. 对每一对数据记录其差值 $X_j - Y_j$ 的符号, 即令

$$z_j = \text{sign}(X_j - Y_j) = \begin{cases} +1, & X_j - Y_j \geqslant 0, \\ -1, & X_j - Y_j < 0. \end{cases}$$

待检验的原假设是

$$H_0 : X_p = Y_p, \tag{12.8.6}$$

X_p, Y_p 分别是总体 X, Y 的 p 分位数. 若原假设为真, 则 z_j=1 和 $z_j = -1$ 的个数应当十分接近; 如果 z_j=1 的个数比 $z_j = -1$ 的个数明显地多或者明显地少, 那么 X_p, Y_p 相差较大, 应当否定原假设. 令 r 为 $z_j = -1$ 的个数, 则它可以作为零假设 $H_0 : X_p = Y_p$ 的检验统计量. 显然, r 是服从二项分布的随机变量, 特别当 H_0 为真, r 服从 p=1/2 的二项分布 $r \sim B(r;\ n,\ p = 1/2)$. 因此, 给定了显著性水平 α, 可利用二项分布的性质确定临界域.

临界域的确定依赖于备择假设 H_1 的选择. 对于 $H_1 : X_p > Y_p$ 的情形, $X_j - Y_j$ 倾向于大于零, 即 z_j 取 +1 的个数较多, r 倾向于取较小的数值; 相反, 对 $H_1 : X_p < Y_p$ 的情形, r 倾向于取大的数值. 因此, 对

(a) $H_0 : X_p = Y_p, H_1 : X_p > Y_p$;

(b) $H_0 : X_p = Y_p, H_1 : X_p < Y_p$;

(c) $H_0 : X_p = Y_p, H_1 : X_p \neq Y_p$

的两子样 p 分位数检验问题, 其临界域可由式 (12.8.2)∼ 式 (12.8.4) 分别决定, 只需将其中的 X_p^0 改成 Y_p, p_0 改成 1/2 即可.

如 (1) 中所述, 当子样容量 n 大时, 可利用正态近似确定临界域.

该方法还可用来检验原假设 $H_0 : X_p - Y_p = C, C$ 为常数, 这时, 取 $z_j = \text{sign}\,(X_j - Y_j - C)$, 其余步骤完全相同.

例 12.15　汽车发动机耗油量 (1)

为了比较两种型号的汽车发动机的油耗量, 将它们各安装在 13 辆同型号的汽车上, 每辆汽车安装两种型号发动机后由同一驾驶员驾驶, 行驶后获得每升油行车公里数的数据如表 12.15 所示.

表 12.15　两种发动机每升油行车公里数

$1^{\#}$ 发动机	4.27	4.75	6.63	7.03	6.78	4.55	5.71	6.00	7.44	4.95	6.16	5.20	4.88
$2^{\#}$ 发动机	4.14	4.90	6.22	6.90	6.87	4.41	5.79	5.80	6.91	4.70	6.20	4.90	4.91

设 μ_1, μ_2 表示装备 $1^{\#}$、$2^{\#}$ 发动机的汽车每消耗 1 升油行驶的平均公里数, 现在检验的原假设和备择假设是

$$H_0 : \mu_1 = \mu_2, H_1 : \mu_1 > \mu_2.$$

子样容量 n 为 13. 其中 $z = -1$ 的个数 $r_{\text{obs}} = 5$. 取水平 $\alpha = 0.05$, 根据式 (12.8.2), 从附表 2 查得 $r_\alpha < 4$, 临界域为 $r \leqslant r_\alpha$ 的区域. 所以 r_{obs} 落在临界域外, 应接受原假设而拒绝备择假设 H_1, 即 $1^{\#}$ 型发动机并不比 $2^{\#}$ 型省油.

(3) 中位数检验法

两子样的符号检验中, 从两个总体中抽取的子样容量必须相同. 在两子样容量不等的情形下, 只能利用容量小的子样观测值构成检验统计量, 一则浪费了测量数

据提供的信息, 二则从容量大的子样中去除若干个观测值可能引起人为的偏差而影响检验的客观性和正确性. 在这种情形, 利用中位数检验法是比较合适的.

设 $X_1, \cdots, X_m$ 和 $Y_1, \cdots, Y_n$ 分别是从总体 X, Y 独立地抽取的随机子样, X 和 Y 的累积分布是连续函数 $F(x)$ 和 $G(y)$. 待检验的原假设是

$$H_0 : \mu_x = \mu_y. \tag{12.8.7}$$

μ_x, μ_y 分别是总体 X, Y 的中位数.

将 $X_1, \cdots, X_m$ 和 $Y_1, \cdots, Y_n$ 按数值递增的次序排列, 构成总体 X 和 Y 的混合子样, 记为

$$z_1 \leqslant z_2 \leqslant \cdots \leqslant z_{m+n}.$$

令 l 是满足下式的正整数：

$$l = \begin{cases} \dfrac{m+n}{2}, & m+n = \text{偶数}, \\ \dfrac{m+n+1}{2}, & m+n = \text{奇数}, \end{cases} \tag{12.8.8}$$

这时, z_l 是混合子样的子样中位数.

令 r 是 X 的子样 $X_1, \cdots, X_m$ 中数值小于等于中位数 z_l 的个数. 如果原假设成立, 即 $\mu_x = \mu_y$, 那么 r 应当接近于 $m/2$. 若 r 接近于 m, 则总体 X 的中位数比 Y 的中位数要小, 倾向于拒绝原假设 $H_0 : \mu_x = \mu_y$ 而接受备择假设 $H_1 : \mu_x < \mu_y$; 相反, 当 r 接近于 0, 则倾向于接受备择假设 $H_1 : \mu_x > \mu_y$. 因此, r 可作为原假设 H_0 的检验统计量.

随机变量 r 的分布为[14]

$$P_{H_0}(r) = \frac{\begin{pmatrix} m \\ r \end{pmatrix} \begin{pmatrix} n \\ l-r \end{pmatrix}}{\begin{pmatrix} m+n \\ l \end{pmatrix}}, \quad r = 0, 1, \cdots, m. \tag{12.8.9}$$

根据以上讨论, 可得到临界域如下：

(a) $H_0 : \mu_x = \mu_y, H_1 : \mu_x > \mu_y$, 临界域 $r \leqslant r_\alpha, r_\alpha$ 满足

$$P_{H_0}(r \leqslant r_\alpha) \leqslant \alpha. \tag{12.8.10}$$

(b) $H_0 : \mu_x = \mu_y, H_1 : \mu_x < \mu_y$, 临界域 $r \geqslant r_\alpha, r_\alpha$ 满足

$$P_{H_0}(r \geqslant r_\alpha) \leqslant \alpha. \tag{12.8.11}$$

(c) $H_0:\mu_x=\mu_y, H_1:\mu_x\neq\mu_y$, 临界域 $r\leqslant r_1$ 和 $r\geqslant r_2$ 满足

$$P_{H_0}(r\leqslant r_1)\leqslant \alpha/2,\quad P_{H_0}(r\geqslant r_2)\leqslant \alpha/2. \tag{12.8.12}$$

作为中位数检验法的一个应用, 我们来考察两种牌号的电池寿命问题. 牌号 A 和 B 的电池寿命 (小时) 有如下数据：

$$\begin{aligned} A:&\ \ 40,\ 30,\ 40,\ 45,\ 55,\ 30;\\ B:&\ \ 50,\ 50,\ 45,\ 55,\ 60,\ 40. \end{aligned}$$

问它们的平均寿命是否可认为相同.

在本问题中, 原假设和备择假设是

$$H_0:\mu_A=\mu_B,\quad H_1:\mu_A<\mu_B.$$

混合子样序列是

$30\leqslant 30\leqslant 40\leqslant 40\leqslant 40\leqslant 45\leqslant 45\leqslant 50\leqslant 50\leqslant 55\leqslant 55\leqslant 60, m=n=6,\ l=(m+n)/2=6$. 混合子样中值是 45, $r_{\text{obs}}=5$. 选定 $\alpha=0.05$, 因为

$$\begin{aligned} P_{H_0}(r\geqslant r_{\text{obs}}) &= \sum_{r=5}^{6} P_{H_0}(r)\\ &= \frac{\begin{pmatrix}6\\5\end{pmatrix}\begin{pmatrix}6\\1\end{pmatrix}}{\begin{pmatrix}12\\6\end{pmatrix}}+\frac{\begin{pmatrix}6\\6\end{pmatrix}\begin{pmatrix}6\\0\end{pmatrix}}{\begin{pmatrix}12\\6\end{pmatrix}}\approx 0.04, \end{aligned}$$

由式 (12.8.11) 可知, r_{obs} 落在临界域内, 故在水平 $\alpha=0.05$ 上拒绝原假设, 而接受牌号 A 的电池寿命短于 B 的备择假设.

12.8.2 两子样的游程检验

设从两个连续总体 X 和 Y 得到容量分别为 n 和 m 的随机子样, 按照观测值增加的次序排列得到子样的顺序统计量 $X_1,\cdots,X_n$ 和 $Y_1,\cdots,Y_m$. 不失一般性, 可假定 $n\leqslant m$(当 $n>m$ 时可交换随机变量 X 和 Y 的记号). 现在的问题是要根据这两组顺序统计量来检验总体 X 和 Y 是否有相同的分布. 于是检验的零假设为

$$H_0: f(x)=f(y). \tag{12.8.13}$$

将 X 和 Y 各自的顺序统计量合并为子样观测值顺序增加的**混合顺序统计量**, 其中共含 $n+m$ 个观测值. 例如, 它可以是这样的一串数值：

$$X_1X_2Y_1X_3Y_2X_4X_5X_6Y_3Y_4Y_5X_7. \tag{12.8.14}$$

如果变量 X 与 Y 具有相同的分布, 即 H_0 为真, 那么 X_i 和 Y_i 可以看成是同一总体中抽取的子样观测值, 因此, 在混合顺序统计量中, X 的观测 X_i 和 Y 的观测 Y_i 应该是充分混合的, 即 X_i 和 Y_i 在混合顺序统计量中出现的位置是比较均匀的; 相反, 若 X_i(或 Y_i) 在混合顺序统计量中集中地出现, 则指示出这两组数据之间存在系统的差异, X_i 和 Y_i 来源于相同总体的可能性很小.

如果在混合顺序统计量中, 将 X 的观测 X_i 用 0 代替, Y 的观测 Y_i 用 1 代替, 就得到一个由 0 和 1 两种元素组成的序列

$$u_1, u_2, u_3, \cdots, u_{n+m}. \quad u_i = \begin{cases} 0, & \text{观测值为} X_i, \\ 1, & \text{观测值为} Y_i. \end{cases} \tag{12.8.15}$$

如果

$$u_{j-1} \neq u_j = u_{j+1} = \cdots = u_{j+l-1} \neq u_{j+l}, \tag{12.8.16}$$

则称 $u_j, \cdots, u_{j+l-1}$ 构成一个**游程**. 式 (12.8.16) 中若 j=1, 则左边的不等号 ($u_{j-1} \neq$) 是多余的; 若 $j + l$−1=$m + n$, 则右边的不等号 ($\neq u_{j+l}$) 是多余的. 换句话说, 一串位置相连、数值相同的元素构成一个游程, 其中元素的个数称为**游程长度**. 在由 0, 1 两种元素构成的序列中, 只有 0 游程和 1 游程两种类型. 例如, 式 (12.8.14) 的混合顺序统计量可表示为

$$\underbrace{00} \quad 10 \quad \underbrace{1000} \quad \underbrace{1110};$$

其中共有 7 个游程, 4 个 0 游程 (2 个长度为 1, 长度为 2, 3 的游程各 1 个); 3 个 1 游程 (2 个长度为 1, 另一个长度为 3). 一个混合顺序统计量数字序列的总游程数用 R 表示, 它取决于混合顺序统计量中 X_i 和 Y_i 的散布状况以及子样容量 n、m, 所以 R 是随机变量.

从前面的讨论可以看到, 当零假设 H_0 不为真时, X_i、Y_i 分属不同的总体, X_i(以及 Y_i) 在混合顺序统计量中倾向于集中出现, 相应地总游程数 R 值就比较小; 而当 H_0 为真, X_i 和 Y_i 均匀地散布于混合顺序统计量中, R 值就比较大. 因此, 总游程数 R 起到指示总体 X 和 Y 之间差异的作用, 可以作为两个总体一致性检验的统计量.

为了利用 R 对两个总体的一致性作显著性检验, 必须求得 H_0 为真时 R 的概率分布. 混合顺序统计量数列式 (12.8.14) 共有 $n + m$ 个元素, 共有 $(n + m)!$ 种排列方式. 因为在 X 的顺序统计量中, n 个 X_i 的相互次序已经固定; 同样 m 个 Y_i 的顺序也已固定, 这使排列方式减少到

$$(n + m)!/n!m! = \begin{pmatrix} m + n \\ n \end{pmatrix}$$

种. 当 H_0 为真, 这 $\begin{pmatrix} m+n \\ n \end{pmatrix}$ 种排列方式出现的概率是相同的.

为了找到游程数等于 R 的概率, 只需要在这些排列方式中找出游程数等于 R 的排列的个数就可以了. 通过直接计算求得游程数 R 的概率分布由下式给定：

$$\begin{aligned} &2 \leqslant R \leqslant 2\min(n,m)+1, \\ &P(R=2k)=\frac{2\begin{pmatrix} n-1 \\ k-1 \end{pmatrix}\begin{pmatrix} m-1 \\ k-1 \end{pmatrix}}{\begin{pmatrix} n+m \\ n \end{pmatrix}}, \\ &P(R=2k+1)=\frac{\begin{pmatrix} n-1 \\ k-1 \end{pmatrix}\begin{pmatrix} m-1 \\ k \end{pmatrix}+\begin{pmatrix} n-1 \\ k \end{pmatrix}\begin{pmatrix} m-1 \\ k-1 \end{pmatrix}}{\begin{pmatrix} n+m \\ n \end{pmatrix}}, \\ &k=1,2,\cdots,\min(n,m). \end{aligned} \tag{12.8.17}$$

其中, 第一个式子表示游程数 R 的最小值为 $2(X_i$ 全部集中于混合顺序统计量的一端, Y_i 全部集中于另一端), 最大值由总体 X 和 Y 中子样容量较小的容量值 $\min(n,m)$ 决定. 游程数分布的均值和方差是

$$E(R)=1+\frac{2nm}{n+m}, \quad V(R)=\frac{2nm(2nm-n-m)}{(n+m)^2(n+m-1)}\ . \tag{12.8.18}$$

对于 $m,n\to\infty$ 的极限情形, 变量

$$Z\equiv\frac{R-E(R)}{\sqrt{V(R)}} \tag{12.8.19}$$

趋近于标准正态分布 $N(0,1)$. 因此, 当 n、m 充分大, 如 m, n >10, 可近似地利用 Z 作为检验统计量; 否则需利用 R 作为检验统计量.

给定显著性水平 α, 对于小样问题, 临界值 R_α 取为满足下述不等式的整数：

$$\sum_{R=2}^{R_\alpha}P(R)\leqslant\alpha<\sum_{R=2}^{R_\alpha+1}P(R). \tag{12.8.20}$$

书末的附表 17 给出了 n, $m=2\to 20$, 四种不同的显著性水平 α 的游程数临界值 R_α. 对于大样问题, 临界值 Z_α 由标准正态函数的积分确定

$$\alpha=\int_{-\infty}^{Z_\alpha}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\mathrm{d}x=\varPhi(Z_\alpha).$$

临界域应为临界值下侧的区域, 因为前面已经阐明, 当 H_0 不为真, 游程数有偏小的趋向. 当根据观测值求得的游程数 $R_{\text{obs}} \leqslant R_\alpha$(或 $Z_{\text{obs}} \leqslant Z_\alpha$), 则在显著性水平 α 上拒绝零假设 H_0, 其中

$$Z_{\text{obs}} = \frac{R_{\text{obs}} - E(R)}{\sqrt{V(R)}}. \tag{12.8.21}$$

故 $Z_{\text{obs}} \leqslant Z_\alpha$ 等价于

$$\Phi(Z_{\text{obs}}) \leqslant \Phi(Z_\alpha) = \alpha. \tag{12.8.22}$$

应当指出, 在两个子样容量 n, m 的数值相差很大的情形下, 游程检验对于两总体一致性的检验效能比较差. 因为在这种情形下, 在混合顺序统计量中, 小容量子样的观测有很大的可能性被大容量子样的观测所隔开, 使得游程数倾向于变大(倾向于两总体一致性的零假设被接受), 但这种状况与两个子样是否属于同一总体却很少有关联. 因此, 游程检验应在 n、m 比较接近时使用. 即使如此, 由于它只利用了数据的很少一部分信息, 故游程检验对于两子样一致性检验的效能不够强. 有时, 游程检验不能排除一致性假设 H_0, 而利用数据信息更充分的其他检验 (如后面将要介绍的斯米尔诺夫检验和威尔科克森检验) 却可能排除 H_0.

例 12.16 两组不变质量测量值之间的一致性

π^0 介子的主要衰变方式是衰变为两个光子

$$\pi^0 \to \gamma\gamma.$$

测量 $\gamma\gamma$ 对的 “不变质量”(见例 3.2 中不变质量的定义)$M_{\gamma\gamma}$ 可以鉴别 π^0 的存在. 两个实验室对产生 π^0 的相同反应中测量了 $M_{\gamma\gamma}$ 值, 得到的顺序测量值 (单位 MeV) 如下所示:

实验 A 81, 82, 87, 93, 102, 104, 108, 112, 116, 122, 125,
($n = 28$) 131, 131, 133, 134, 139, 139, 142, 144, 146, 152,
156, 182, 202, 206, 216, 226, 270.

实验 B 8, 12, 14, 16, 22, 26, 26, 50, 64, 68, 76, 79, 83,
($m = 32$) 88, 96, 97, 98, 99, 103, 105, 107, 113, 114, 115,
126, 128, 130, 132, 138, 150, 169, 171.

要求检验两个实验的结果是否一致.

两组数据的混合顺序测量值可列出数列:

$$b_1\ b_2\ b_3\ b_4\ b_5\ b_6\ b_7\ b_8\ b_9\ b_{10}\ b_{11}\ b_{12}\ a_1\ a_2\ b_{13}\ a_3\ b_{14}\ a_4\ b_{15}$$
$$b_{16}\ b_{17}\ b_{18}\ a_5\ b_{19}\ a_6\ b_{20}\ b_{21}\ a_7\ a_8\ b_{22}\ b_{23}\ b_{24}\ a_9\ a_{10}\ a_{11}\ b_{25}\ b_{26}$$
$$b_{27}\ a_{12}\ a_{13}\ b_{28}\ a_{14}\ a_{15}\ b_{29}\ a_{16}\ a_{17}\ a_{18}\ a_{19}\ a_{20}\ b_{30}\ a_{21}\ a_{22}\ b_{31}\ b_{32}$$
$$a_{23}\ a_{24}\ a_{25}\ a_{26}\ a_{27}\ a_{28}.$$

该数列的游程数 $R_{\text{obs}} = 24$. 因为两子样容量 n、m 很大, 可以利用大样近似公式 (12.8.21), 式 (12.8.22). 游程数 R 的期望值和方差由式 (12.8.18) 求得

$$E(R) = 1 + \frac{2 \times 28 \times 32}{28 + 32} \approx 30.87.$$

$$V(R) = \frac{2 \times 28 \times 32\,(2 \times 28 \times 32 - 28 - 32)}{(28+32)^2 \cdot (28 + 32 - 1)} \approx 14.62.$$

Z 的观测值

$$Z_{\text{obs}} = \frac{R_{\text{obs}} - E(R)}{\sqrt{V(R)}} \approx -1.80.$$

查累积标准正态函数表, 得

$$\varPhi(Z_{\text{obs}}) = \varPhi(-1.80) = 1 - \varPhi(1.80) \approx 0.036.$$

如果给定显著性水平 $\alpha = 0.05$, 则这两组数据一致性的假设被拒绝.

观察两组测量值可以发现, 它们之间最明显的差别在于实验室 B 的数据中有许多数值很小的 $M_{\gamma\gamma}$ 值. 这些小 $M_{\gamma\gamma}$ 值的事例可能是实验 B 的系统偏差造成的. 例如, π^0 衰变成两个 γ 光子, 一个 γ 转化为 e^+e^-, 其中一个 e 发生韧致辐射, 实验中错误地将韧致辐射光子与另一个 γ 合并在一起求出 $M_{\gamma\gamma}$ 值, 它的数值就一定偏小.

假如从 B 实验的数据中除去 7 个 $M_{\gamma\gamma} < 40$ 的 “错误” 事例, 尔后应用类似的步骤, 这样, 子样容量变成 28 和 25, 其结果为 $\varPhi(Z_{\text{obs}}) \approx 0.17$. 即使在 $\alpha = 0.10$ 显著性水平上, 两组观测值也可视为一致.

12.8.3　游程检验作为皮尔逊 χ^2 检验的补充

在拟合优度的皮尔逊 χ^2 检验中, 检验统计量

$$X^2 \equiv \sum_{i=1}^{N} \frac{(n_i - np_i)^2}{np_i}$$

是实验频数与零假设的理论频数之差 (离差) 的平方和. 在这一检验统计量中, 离差的符号, 正负离差出现的先后次序和排列状况不起作用, 数据的这些信息在皮尔逊 χ^2 检验中被损失了. 而游程检验正好利用了数据的这些信息, 因而两者联合使用时, 对拟合优度的检验更为有效.

考虑图 12.12 中所示的三种情况, 其中平滑曲线表示待检验的零假设预期的分布, 直方图为观测到的分布. 假定直方图有 N 个小区间, 定义变量

$$u_i = \begin{cases} 0, & \text{若}_i - np_i < 0 \\ 1, & \text{若}_i - np_i \geqslant 0 \end{cases}, \qquad i = 1,\ 2,\ \cdots,\ N. \tag{12.8.23}$$

由此得到 0, 1 两种元素组成的序列

$$u_1,\ u_2,\ \cdots,\ u_N. \tag{12.8.24}$$

在图 12.12 (a) 的情形下, 观测的分布与零假设 H_0 的分布十分接近, 理论频数与观测频数也十分接近, 离差的符号经常变换, 正离差和负离差的散布是均匀而随机的, 序列 $u_1,\ u_2,\ \cdots,\ u_N$ 的游程数将比较大. 如果零假设的理论分布与观测到的分布虽然形状相近, 但位置上有明显的偏移, 如图 12.12 (b) 所示, 则将出现一连串符号相同的离差; 若零假设的理论分布与测定的分布在形状上不同, 如图 12.12 (c) 所示, 也将出现一连串符号相同的离差. 这两种情况都将导致游程数减小. 所以序列式 (12.8.24) 的游程数 R 可作为零假设 H_0 是否正确的检验统计量.

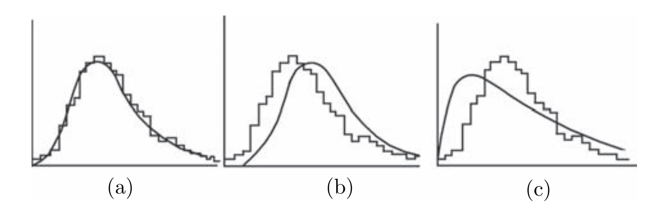


图 12.12　零假设的分布与观测到的分布之比较

(a) 两者的形状和位置相符合; (b) 形状符合, 但位置有偏移; (c) 形状不一致

在序列 $u_1,\ u_2,\ \cdots,\ u_N$ 中, $u_i = 0$ 的个数记为 n, $u_i = 1$ 的个数记为 m, 则游程数 R 的概率分布由式 (12.8.17) 表示; 当 n、m 充分大, 则由式 (12.8.19) 表示. 显然, 这时公式中的 $n+m$ 应由 N 代替. 利用 12.8.2 节中描述的步骤, 可以决定是否应当在给定的显著性水平 α 上排除零假设 H_0.

可以证明, 当 H_0 为简单零假设时, 游程检验与 χ^2 检验是渐近地相互独立的; 如果 H_0 中包含未知参数, 其数值要从观测数据来估计, 则两种检验互相不独立 [39]. 在后一种情况下, 两种检验同时使用的意义不大; 如果两者相互独立, 则能互为补充; 事实上, 它们可以合并为一个单独的检验. 令 P_1 是皮尔逊 χ^2 检验中统计量 X^2 大于观测值 X^2_{obs} 的概率, P_2 是游程检验中游程数 R 小于观测值 R_{obs} 的概率, 当子样容量 N 充分大, P_1 和 P_2 近似地可视为 [0, 1] 区间内均匀分布的随机变量. 这时, 有

$$U = -2\left(\ln P_1 + \ln P_2\right) \sim \chi^2(4). \tag{12.8.25}$$

证明如下：设 x 是 [0, 1] 区间内均匀分布的随机变量, 定义 $u \equiv -2\ln x$, 则 u 的概率密度可表示为 (见式 (2.3.3))

$$g(u) = f(x)\left|\frac{\mathrm{d}x}{\mathrm{d}u}\right| = \frac{1}{2}\exp\left(-\frac{u}{2}\right).$$

这正是自由度为 2 的 χ^2 分布的概率密度, 即 $u \sim \chi^2(2)$. 若 $u_1, u_2, \cdots, u_r$ 是相互独立的 $\chi^2(2)$ 变量, 由 χ^2 分布自由度的可加性知

$$\sum_{i=1}^{r} u_i = -2\sum_{i=1}^{r} \ln x_i \sim \chi^2(2r),$$

证毕.

因此, 式 (12.8.25) 定义的随机变量 U 可作为拟合优度的皮尔逊 χ^2 检验和游程数检验的统一检验统计量. 对于给定的显著性水平 α, 临界值 $\chi^2_\alpha(4)$ 由 $\chi^2(4)$ 的概率密度 $f(y)$ 的积分求得：

$$\alpha = \int_{\chi^2_\alpha(4)}^{\infty} f(y)\mathrm{d}y. \tag{12.8.26}$$

当 U 的观测值 $U_{\text{obs}} > \chi^2_\alpha(4)$, 则在水平 α 上排除零假设 H_0.

例 12.17　实验直方图与理论分布的比较

图 12.13 是粒子反应

$$\bar{\mathrm{p}} + \mathrm{p} \to \pi^+ + \pi^+ + \pi^- + \pi^-$$

中靶质子 p 到 π^- 的四动量转移平方 (用符号 t 表示) 的分布, 其中直方图是测量数据, 平滑曲线是零假设 H_0 的理论分布, 入射反质子 $\bar{\mathrm{p}}$ 的动量为 1.2GeV/c. 曲线下的面积和直方图下的面积 (测量到的总事例数) 同为 $n = 990$. 直观地审视图 12.13 可以得到一个印象：观测到的分布在 $t = 0$ 附近比理论分布有更多的事例.

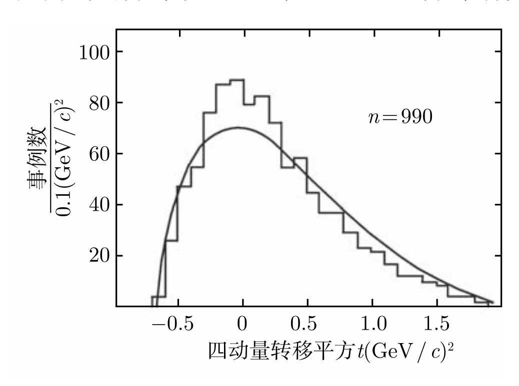


图 12.13　四动量转移平方的分布

为了定量地检验零假设, 首先用皮尔逊 χ^2 检验方法. 为了满足每个子区间中事例数不能太少这一要求, 将 $t > 1.6(\mathrm{GeV}/c)^2$ 的数据归并到一个子区间中, 这样, 共计有 24 个子区间, 只有第一个子区间内理论频数小于 5, 符合 χ^2 检验的要求.

每个子区间内的理论频数 np_i 可由曲线的数值积分求出, 计算的结果为

$$X_{\text{obs}}^2=\sum_{i=1}^{24}\frac{(n_i-np_i)^2}{np_i}=\sum_{i=1}^{24}\frac{n_i^2}{np_i}-n\approx 30.6.$$

由于对理论分布归一化条件的存在, χ^2 检验的自由度 $\nu=24-1=23$. 观测值 X_{obs}^2 对应的积分 χ^2 概率 $P_{\chi^2}=P_1=0.14$(见式 (12.4.23) 的定义). 因此, χ^2 检验的结果表明, 在通常选定的显著性水平上 ($\alpha=0.05-0.10$), 没有理由拒绝 H_0.

现在再对同样的数据做游程检验. 从图 12.13 可见, 在 7 个子区间中观测值高于理论曲线, 17 个子区间中则恰恰相反, 游程数的观测值 $R_{\text{obs}}=5$. 将 $n=7,m=17,R\leqslant R_{\text{obs}}=5$ 代入式 (12.8.17), 求得在 7+17 个元素序列中出现不多于 5 个游程的积分概率为

$$P_2=0.0034.$$

从附表 17 查到, 对于 $\alpha=0.05\sim 0.10$, 相应的临界值 $R_\alpha>6$, 所以游程检验倾向于排除零假设 H_0.

在本例中, H_0 是简单原假设, 理论分布没有待定的未知参数, 故可以做联合检验. 联合检验统计量 U 的观测值是

$$U_{\text{obs}}=-2(\ln P_1+\ln P_2)\approx 15.3.$$

因为 $U\sim\chi^2(4)$, 所以 U_{obs} 对应的积分 χ^2 概率

$$P_{U_{\text{obs}}}\lesssim 0.005,$$

它远小于通常选定的显著性水平 $\alpha=0.05\sim 0.10$, 所以应当拒绝 H_0. 该结论与游程检验的结果一致而与 χ^2 检验的结果相反, 这说明本例中游程检验作为 χ^2 检验的补充对零假设 H_0 的检验是完全必要的.

12.8.4 两子样的斯米尔诺夫检验

在许多实际问题中, 常常希望根据两个总体各自的一组子样测定值来确定这两个总体是否具有相同的分布. 斯米尔诺夫 (Smirnov) 检验利用两个总体的子样分布函数给出了与柯尔莫哥洛夫检验相类似的检验统计量. 斯米尔诺夫检验适用于连续总体的检验问题.

设总体 X 和 Y 的累积分布函数分别是 $F(x)$ 和 $G(x)$, 则两总体一致性检验的零假设和备择假设为

$$\begin{cases}H_0:F(x)=G(x),H_1:F(x)\neq G(x),\\ H_0:F(x)=G(x),H_1:F(x)>G(x),\\ H_0:F(x)=G(x),H_1:F(x)<G(x).\end{cases}\tag{12.8.27}$$

设 $X_1,\cdots,X_m$ 和 $Y_1,\cdots,Y_l$ 分别是总体 X 和 Y 的子样, $S_m(x)$ 和 $S_l(x)$ 分别是它们的子样分布函数. 统计量 $D_{m,l}$, $D^+_{m,l}$, $D^-_{m,l}$ 定义为

$$\begin{cases} D^+_{m,l} = \max\limits_{-\infty<x<\infty}[S_m(x)-S_l(x)], \\ D^-_{m,l} = \max\limits_{-\infty<x<\infty}[S_l(x)-S_m(x)], \\ D_{m,l} = \max\limits_{-\infty<x<\infty}|S_m(x)-S_l(x)| = \max(D^+_{m,l}, D^-_{m,l}). \end{cases} \tag{12.8.28}$$

参考 12.4.5 节的讨论可知, 这三个统计量反映了 $F(x)$ 和 $G(x)$ 之间的差异, 可作为两个总体一致性的检验统计量.

当 H_0 为真, 且 $F(x)$, $G(x)$ 为连续函数, 子样容量 $m=l$, 则 $D^+_{m,m}$ 和 $D_{m,m}$ 的概率分布为 ($D^+_{m,l}$ 与 $D^-_{m,l}$ 有相同分布)

$$P(D^+_{m,m}<z) = \begin{cases} 0, & 当 z\leqslant 0, \\ 1, & 当 z>1, \\ 1-\dfrac{\begin{pmatrix} 2m \\ m-c \end{pmatrix}}{\begin{pmatrix} 2m \\ m \end{pmatrix}}, & 当 0<z\leqslant 1; \end{cases}$$

$$P(D_{m,m}<z) = \begin{cases} 0, & 当 z<\frac{1}{m}, \\ 1, & 当 z>1, \\ \sum\limits_{j=-\left[\frac{m}{c}\right]}^{\left[\frac{m}{c}\right]}(-1)^j\dfrac{\begin{pmatrix} 2m \\ m-j \end{pmatrix}}{\begin{pmatrix} 2m \\ m \end{pmatrix}}, & 当 \dfrac{1}{m}<z\leqslant 1, \end{cases}$$

其中 $c=-[-zm]$, 方括号表示取整数.

当 $m\neq l$, 但 $m,l\to\infty$ 的极限情形下, 则有

$$\lim_{\substack{m\to\infty \\ l\to\infty}} P\left(D^+_{m,l}\leqslant\frac{z}{\sqrt{n}}\right) = 1-\mathrm{e}^{-2z^2}, \quad z>0$$

$$\lim_{\substack{m\to\infty \\ l\to\infty}} P\left(D_{m,\,l}\leqslant\frac{z}{\sqrt{n}}\right) = 1-2\sum_{i=1}^{\infty}(-1)^{i-1}\mathrm{e}^{-2i^2z^2}, \quad z>0 \tag{12.8.29}$$

其中,

$$n=\frac{ml}{m+l}.$$

$D_{m,l}^{+}$、$D_{m,l}$ 的极限分布与柯尔莫哥洛夫检验统计量 D_n^+、D_n 的极限分布形式相同. 类似地, 由 $D_{m,\,l}^{+}$ 极限分布的表达式, 立即有 $(m, l \to \infty)$

$$D_{m,l}^{+} \sim \sqrt{\frac{\chi^2(2)}{4n}}. \tag{12.8.30}$$

由以上概率分布可知, 斯米尔诺夫检验是分布自由的, 适用于任何连续总体.

对于不同的备择假设, 适当的检验统计量如下：

原假设	备择假设	检验统计量
$H_0 : F(x) = G(x)$	$H_1 : F(x) \neq G(x)$	$D_{m,l}$
$H_0 : F(x) = G(x)$	$H_1 : F(x) > G(x)$	$D_{m,l}^{+}$
$H_0 : F(x) = G(x)$	$H_1 : F(x) < G(x)$	$D_{m,l}^{-}$

检验的方法与柯尔莫哥洛夫检验完全相似. 给定显著性水平 α, 大于临界值 $D_{m,l;\alpha}^{(+,-)}$ 的区域为临界域. 对于五种不同的显著性水平 α 值, 附表 18 给出临界值 $D_{m,m;\alpha}^{(+)}$, 附表 19 给出临界值 $D_{m,l;\alpha}^{(+)}$. 两个表的最后一行都给出了 $m, l \to \infty$ 极限情形下的临界值表达式. 对于单侧备择假设的情形, 当 m、l 很大时, 还可利用式 (12.8.30) 确定临界值

$$D_{m,l;\alpha}^{+,-} = \sqrt{\frac{\chi_\alpha^2(2)}{4n}}, \quad n = \frac{ml}{m+l}.$$

$\chi_\alpha^2(2)$ 是 $\chi^2(2)$ 分布的上侧 α 分位数.

作为斯米尔诺夫检验应用的一个例子, 我们回过头来讨论例 12.16 两组不变质量测量值之间的一致性检验问题. 两个实验室测定值的分布 S_{28}(实验 A) 和 S_{32}(实验 B) 示于图 12.14. 两个阶梯函数的最大差值是

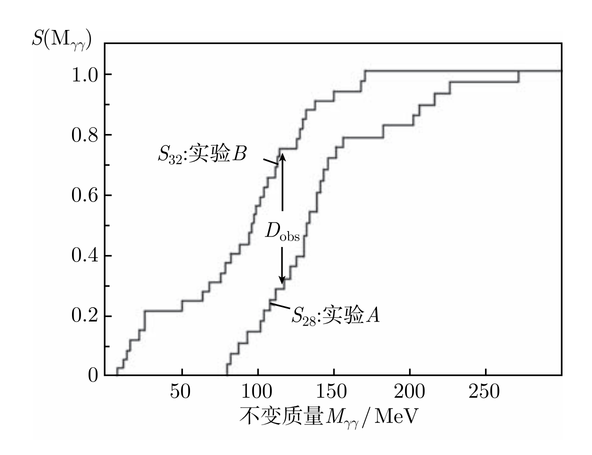


图 12.14　例 12.16 中两组不变质量测量值的累积分布

$$D_{\text{obs}} = \max|S_{32}(M_{\gamma\gamma}) - S_{28}(M_{\gamma\gamma})| = \frac{24}{32} - \frac{8}{28} = 0.46,$$

选定显著性水平 $\alpha = 0.05$. 从附表 19 查得 $D_{28,32;0.05}$ 的值为 0.352. 因此, D_{obs} 落在临界域内, 在显著性水平 $\alpha = 0.05$ 上必须拒绝两个子样一致性的零假设, 这与游程检验的结论一致.

假如像例 12.16 中所述, 实验 B 可能有系统的偏差, $M_{\gamma\gamma}$ 过小的 7 个事例可能是错误的而必须舍弃, 这样, 实验 B 的子样容量变成 25. 游程检验的结果是, 对于修正后的两组测量值, 即使在 $\alpha = 0.10$ 的显著性水平上也可视为一致. 现在使用斯米尔诺夫检验, 对于修正后的两组测量值

$$D_{\text{obs}} = \max|S_{25}(M_{\gamma\gamma}) - S_{28}(M_{\gamma\gamma})| = \frac{17}{25} - \frac{8}{28} = 0.39.$$

由附表 19 查得 $D_{25,28;0.05} = 0.374$, $D_{25,28;0.10} = 0.336$. 对于显著性水平 $\alpha = 0.05$ 或 $\alpha = 0.10$, D_{obs} 都落在临界域内. 所以斯米尔诺夫检验的结果是应当排除两组测量值 (B 实验的测量数据经过修正) 一致性的零假设, 该结论与游程检验的结果矛盾. 由于游程检验对数据信息的利用不充分, 我们宁可采用斯米尔诺夫检验的结果.

12.8.5 两子样的威尔科克森检验

(1) 分布函数一致性的检验

设两个连续总体各自的子样顺序统计量分别为 $X_1, \cdots, X_n$ 和 $Y_1, \cdots, Y_m$, 不失一般性可假定 $n \leqslant m$. 待检验的零假设是这两个子样产生于同样的总体.

将这 $n+m$ 个观测按数值的递增次序排列为混合顺序统计量, 每一个观测值的**秩**定义为该值在混合顺序统计量这个数列中的**顺序号**, 若该数列中有 l 个观测值相等, 则它们的秩都规定为它们的顺序号的平均值. 威尔科克森 (Wilcoxon) 检验中的检验统计量 W 等于子样 $X_1, \cdots, X_n$ 的**秩的总和**, 所以这种检验也称为**秩和检验**.

如 12.8.2 节游程检验中所述, 若 H_0 不为真, 即两子样产生于不同的总体, 则 $X_1, \cdots, X_n$ 与 $Y_1, \cdots, Y_m$ 之间存在着系统的差异. 在混合顺序统计量数列中, X 子样值将比较集中于数列的某一边. 两种极端的情况是 X 子样观测值全部出现于数列左端, 以及全部出现于数列右端. 对于第一种情形, 秩总和达到极小, 即头 n 个正整数之和

$$W_{\min} = n(n+1)/2. \tag{12.8.31}$$

对于第二种情形, 秩总和达到极大

$$W_{\max} = \sum_{j=m+1}^{m+n} j = \frac{n(n+1)}{2} + mn. \tag{12.8.32}$$

当 H_0 为真, X 子样和 Y 子样产生于同一总体, 那么在混合顺序统计量中, 它们将充分均匀地混合出现, 秩总和 W 应当远离 $W_{\min}$ 和 $W_{\max}$ 这两个极端值, 而比较

接近于这两者的平均. 因此, 统计量 W 对于 H_0 是否为真敏感, 可以作为 H_0 的检验统计量. 而且由上述讨论可知, W 接近 $W_{\min}$ 和 $W_{\max}$ 的区域都将是临界域, 故必须作双侧检验.

秩总和 W 在 $W_{\min} \sim W_{\max}$ 的概率分布可按照类似游程检验中游程分布的方法导出. 其结果是 W 对均值为对称分布, 其均值和方差为

$$E(W) \equiv \bar{W} = \frac{n}{2}(n+m+1), \quad V(W) = \frac{nm}{12}(n+m+1). \tag{12.8.33}$$

注意到 $\bar{W}$ 恰好是 $W_{\min}$ 和 $W_{\max}$ 的算术平均.

当子样容量 $m \leqslant 25$, 对于六种不同的显著性水平 α 值, 附表 20 给出了统计量 W 的临界值 W_α. W_α 的定义是满足下式的整数：

$$\sum_{W_{\min}}^{W_\alpha} P(W) \leqslant \alpha < \sum_{W_{\min}}^{W_\alpha+1} P(W). \tag{12.8.34}$$

这样定义的临界值只适合做下侧检验. 为了得到显著性水平 α 的双侧临界值, 可首先从表中查得显著性 $\alpha/2$ 的下侧临界值 W_l; 上侧临界值 W_u 根据分布的对称性算出

$$W_u = \bar{W} + (\bar{W} - W_l) = 2\bar{W} - W_l. \tag{12.8.35}$$

附表 20 中也列出了一定 n、m 对应的 $2\bar{W}$ 值.

当 n、m 很大时, 可以证明 W 的分布趋近于正态分布. 故对大样问题, 威尔科克森检验可利用渐近统计量

$$Z = \frac{W - \bar{W} \pm \frac{1}{2}}{\sqrt{V(W)}} \sim N(0,1). \tag{12.8.36}$$

计算下侧积分概率时, 式 (12.8.36) 分子中取 $+1/2$; 计算上侧积分概率时取 $-1/2$. 对于实际问题, 当 n、$m > 10$, 正态分布便是很好的近似. 倘若 m 比 n 大得不多, 而显著性 α 又不太小 (比如 $\alpha > 0.01$), 即使 $n < 10$, 这一近似仍然可用.

例 12.18 π^0 寿命两组测量值的一致性检验

若干个实验组用两种不同的方法测量 π^0 介子的平均寿命. 一类实验利用核乳胶作为探测装置, 另一类实验则用计数器. 平均寿命的两组测量结果汇总如下 (忽略测量误差)：

计数器测定值 (10^{-16}s)：0.56 0.6 0.73 0.9 1.05

核乳胶测定值 (10^{-16}s)：1.0 1.6 1.7 1.9 2.3 2.8

问在显著性水平 $\alpha = 0.05$ 上两种方法的测量结果是否一致.

应用秩和检验方法来检验两子样总体的一致性零假设. 现在 $n=5, m=6$, 由附表 20 查得, 当 $\alpha=0.05$ 时, 统计量 W 的下侧和上侧临界值是

$$W_l = W_{.025} = 18, \quad W_u = 2\bar{W} - W_l = 60 - 18 = 42.$$

混合顺序统计量的数列是

$$\underline{0.56}\ \ \underline{0.6}\ \ \underline{0.73}\ \ \underline{0.9}\ \ 1.0\ \ \underline{1.05}\ \ 1.6\ \ 1.7\ \ 1.9\ \ 2.3\ \ 2.8$$

其中有横线的数据是计数器实验的测量值. 于是秩总和的观测值为

$$W_{\text{obs}} = 1+2+3+4+6 = 16.$$

显然, $W_{\text{obs}} < W_l$, 即 W_{obs} 落入下侧临界域. 因此, 根据秩和检验, π^0 平均寿命的两组测量值在 $\alpha=5\%$ 显著性水平上是不一致的.

如果对上述数据作游程检验, 则混合顺序统计量数列的游程数观测值 $R_{\text{obs}}=4$. 从附表 17 可查出 $n=5, m=6, \alpha=0.05$ 对应的游程数临界值是 $R_{.05}=3$. 因此, 根据游程检验, 在 $\alpha=0.05$ 显著性水平上不能拒绝两组实验结果一致的零假设, 这与威尔科克森检验的结果不同. 由于后者利用数据的信息比游程检验充分, 它对一致性检验的效能较强, 所以宁可采用秩和检验的结论.

(2) 分位数一致性的检验

威尔科克森方法还可应用于检验两个总体 X 和 Y 的分位数的一致性, 即待检验的原假设为

$$H_0: X_p = Y_p. \tag{12.8.37}$$

与两子样的符号检验 (见 12.8.1 节) 中相同, X 和 Y 的子样必须成对地抽取. 设有 n 对随机抽样值 $(X_1, Y_1), \cdots, (X_n, Y_n)$, 用 z_j 表示 $X_j - Y_j$ 的符号

$$z_j = \text{sign}(X_j - Y_j) = \begin{cases} +1, & X_j - Y_j > 0 \\ -1, & X_j - Y_j < 0. \end{cases} \tag{12.8.38}$$

需要注意, 与符号检验中不同的是对于 $X_j - Y_j = 0$ 的数据在这里必须被舍弃. 记

$$d_j = X_j - Y_j, \quad D_j = |d_j|, \quad j = 1, 2, \cdots, n. \tag{12.8.39}$$

将 n 个 D_j 按数值递增的次序排成一个数列, 每个 D_j 对应的序号被规定为它的秩 R_j, 若序号 $m+1$ 到 $m+l$ 的 l 个 D_j 数值相等, 则它们的秩都规定为序号的平均值 $m+(l+1)/2$.

定义

$$W_- = \sum_{j=1}^{n} |z_j| R_j, \quad 对 z_j = -1,$$

$$W_{+}=\sum_{j=1}^{n}z_{j}R_{j},\quad \text{对}z_{j}=+1, \tag{12.8.40}$$

则 W_{-} 与 W_{+} 分别是 z_j 为负与正的秩的总和. 当 H_0 为真, 两总体分位数 X_p 与 Y_p 相等, 直观地可知 W_{-} 和 W_{+} 应当很接近, 而且有

$$W_{-}\approx W_{+}\approx\frac{1}{2}\sum_{i=1}^{n}R_{j}=\frac{1}{2}\sum_{j=1}^{n}j=\frac{n}{4}(n+1).$$

相反, 若 H_0 不为真, 则 W_{-} 与 W_{+} 的差别增大, 特别对于所有 z_j 都大于 (小于) 零的极端情形, $W_{-}(W_{+})$ 等于它的极大值

$$(W_{-})_{\max}=(W_{+})_{\max}=\sum_{j=1}^{n}R_{j}=\sum_{j=1}^{n}j=\frac{n}{2}(n+1).$$

显然, $X_p=Y_p$ 的原假设不成立. 因此, W_{-} 或 W_{+} 可作为原假设 $H_0: X_p=Y_p$ 的检验统计量. 当 H_0 为真, 由于对称性, W_{-} 和 W_{+} 有相同的分布. 我们利用随机变量

$$W=\min(W_{-},W_{+}) \tag{12.8.41}$$

作为检验统计量, 称为**威尔科克森符号秩和检验统计量**. 只要知道了 W 在 H_0 为真情形下的概率分布, 根据假设检验的一般方法 (见 12.1.2 节), 便可求出显著性水平 α 上的临界域. 对于单侧备择假设

$$H_1: X_p<Y_p,\quad (W=W_{+}), \tag{12.8.42}$$

$$H_1: X_p>Y_p,\quad (W=W_{-}); \tag{12.8.43}$$

临界域为 $W\leqslant W_\alpha, W_\alpha$ 满足

$$P_{H_0}(W\leqslant W_\alpha)<\alpha. \tag{12.8.44}$$

对于双侧备择假设

$$H_1: X_p\neq Y_p,\quad \left[W=\min\left(W_{-},W_{+}\right)\right], \tag{12.8.75}$$

临界域为 $W\leqslant W_\alpha, W_\alpha$ 满足

$$P_{H_0}(W\leqslant W_\alpha)<\alpha/2. \tag{12.8.46}$$

当从 n 对子样观测值 $(x_j,y_j)(j=1,2,\cdots,n)$ 计算出 W 的观测值 W_{obs}, 若 W_{obs} 落在临界域内, 则在显著性水平 α 上拒绝原假设 H_0.

现在来推导 W 的分布. 当 H_0 为真, 则每个 d_j 取正号和取负号的概率均为 1/2, 与 d_j 对应的秩值 z_jR_j 取正号或负号的可能性相等, 也就是有两种等可能的方式

$$z_jR_j, \quad z_j=-1,+1.$$

现在有 n 个差值 d_j, 对应于 n 个秩值, 将他们排成数列

$$z_1R_1, z_2R_2, \cdots, z_nR_n.$$

每个 z_jR_j 有两种等可能的方式 $+R_j$、$-R_j$, 故该数列共有 2^n 种等可能的方式. 设这 2^n 种方式中使得 $W \leqslant W_\alpha$ 成立的方式个数是 $m(W \leqslant W_\alpha)$, 变量 W 的概率分布可表示为

$$P_{H_0}(W \leqslant W_\alpha)=\frac{m(W \leqslant W_\alpha)}{2^n}. \tag{12.8.47}$$

基于这种考虑, 对于子样容量 $5 \leqslant n \leqslant 30$, 附表 21 给出了三种显著性水平 α 的近似临界值 W_α. 在 $n \to \infty$ 的极限情形下, 当 H_0 为真, W 趋近于正态分布, 均值和方差分别为

$$\mu_W=\frac{n(n+1)}{4}, \quad \sigma_W^2=\frac{n(n+1)(2n+1)}{24}. \tag{12.8.48}$$

故对于 $n>30$ 的情形, 可利用性质

$$z=\frac{W-\mu_W}{\sigma_W} \sim N(0,1) \tag{12.8.49}$$

来确定临界值.

上述方法也可用来检验原假设 $H_0: X_p-Y_p=C$(常数), 只需将 d_j 定义为 X_j-Y_j-C 即可; 同样还适用于单子样的分位数 X_p 是否等于某个常数 X_p^0 的检验问题, 此时 d_j 定义为 $X_j-X_p^0$, 其余步骤则完全相同.

由以上讨论可见, 符号检验中只利用了两总体 n 对子样观测值之差的符号信息, 而威尔科克森符号秩和检验则同时利用了 n 对观测值之差的符号和数值, 因此, 后者对于两个总体分位数差别的检验更加灵敏, 结果自然也更为可靠.

例 12.19 汽车发动机耗油量 (2)

现在我们用威尔科克森符号秩和方法来重新检验 12.8.1 节中 (符号检验) 的发动机油耗问题, 对于同样的数据, 采用本节的符号, 如表 12.16 所示.

与 12.8.1 节中相同, 原假设和备择假设是

$$H_0: \mu_x=\mu_y, \quad H_1: \mu_x>\mu_y.$$

表 12.16　两种发动机耗油量的威尔科克森符号秩和检验

X_i	Y_i	d_i	z_i	R_i
4.27	4.14	0.13	1	5.5
4.75	4.90	−0.15	−1	8
6.63	6.22	0.41	1	12
7.03	6.90	0.13	1	5.5
6.78	6.87	−0.09	−1	4
4.55	4.41	0.14	1	7
5.71	5.79	−0.08	−1	3
6.00	5.80	0.20	1	9
7.44	6.91	0.53	1	13
4.95	4.70	0.25	1	10
6.16	6.20	−0.04	−1	2
5.20	4.90	0.30	1	11
4.88	4.91	−0.03	−1	1

显著性水平选为 $\alpha = 0.05$. 由表 12.16 可知

$$W_{\text{obs}} = (W_{-})_{\text{obs}} = 8 + 4 + 3 + 2 + 1 = 18.$$

查阅附表 21, 当子样容量 $n = 13$, $\alpha = 0.05$, 临界值是 $W_\alpha = 21$, 临界域为 $W \leqslant W_\alpha$ 的区域. 因此, W_{obs} 落在临界域内, 应接受备择假设, 即 $1^{\#}$ 发动机比 $2^{\#}$ 发动机省油, 该结论与符号检验的结果不同. 由于威尔科克森方法比符号检验更灵敏, 我们宁可采用前者的结论.

如果我们仔细检查一下数据 X_j、Y_j, 就不难了解两种检验方法得出不同结论的原因. 符号检验只利用 $X_j - Y_j$ 的符号信息 (即 z_j), 13 对数据中 $z_j = -1$ 的数据有 5 对, 与 $z_j = +1$ 的数据 8 对相比相差不大, 因而得出两者均值大体相同的结论. 但在 $z_j = -1$ 的 5 对数据中, $X_j - Y_j$ 之差的数值大多比较小, 因而它们的秩比较小. 将这一因素考虑在内, 威尔科克森方法得到 $\mu_x > \mu_y$ 的结论. 显然后者更为合理. 事实上, 分别对 X_j 和 Y_j 的数据作简单平均得到 $\bar{X} = 5.73$, $\bar{Y} = 5.59$, 两者的差别达到 2.5%.

12.8.6　多个连续总体子样的克鲁斯卡尔–瓦列斯秩检验

当需要对 $J > 2$ 组测量数据做总体一致性检验时, 可以应用威尔科克森的子样检验方法对每两组数组做一致性检验, 这样共需要做 $\dfrac{J}{2}(J-1)$ 次检验. 当 J 不很小时, 检验次数很多, 整个过程很繁琐. 同时, 即使所有各组测量值确实来源于同一总体, 但由于随机涨落的存在, 仍有可能某两组数据被秩和检验判定为不一致, 这就大大增加了排除真实假设的可能性. 最后, 这种 “两两成对” 检验一致性的方法不能给出所有各组测量值之间整体一致性的数值表示. 因此, 简单地将两子样一致

性的检验方法应用于多个子样一致性检验是不适当的.

对任意多个子样作一致性检验的有效方法是克鲁斯卡尔–瓦列斯 (Kruskal-Wallis) 秩检验. 设从 J 个连续总体抽取了各自的随机子样, 总共包含 N 个观测, 按数值递增次序排列成混合顺序统计量数列, 每个观测的秩等于它在该数列中的顺序号 $(1 \to N)$. 第 j 个子样的容量记为 n_j, 它的秩和为 W_j, 平均秩 $\overline{W_j} = W_j/n_j(j = 1, \cdots, J)$. 如果 H_0 为真, 即所有 J 个子样属于同一个总体, 预期这 J 个子样大致上有相同的平均秩和相同的平均秩方差

$$E(\overline{W_j}) = \frac{1}{N}\sum_{i=1}^{N} i = \frac{1}{2}(N+1), \tag{12.8.50}$$

$$V(\overline{W_j}) = \frac{1}{N-1}\sum_{i=1}^{N}\left[i - E(\overline{W_j})\right]^2 = \frac{1}{12}N(N+1). \tag{12.8.51}$$

K-W(即克鲁斯卡尔–瓦列斯) 秩检验使用的检验统计量是

$$H \equiv \sum_{j=1}^{J} n_j(\overline{W_j} - E(\overline{W_j}))^2/V(\overline{W_j}). \tag{12.8.52}$$

写成易于计算的形式有

$$\begin{aligned} H &= \sum_{j=1}^{J} n_j\left(\overline{W_j} - \frac{1}{2}(N+1)\right)^2 \Bigg/ \left(\frac{1}{12}N(N+1)\right) \\ &= \frac{12}{N(N+1)}\sum_{j=1}^{J}\frac{W_j^2}{n_j} - 3(N+1), \end{aligned} \tag{12.8.53}$$

在推算过程中用到关系式

$$\sum_{j=1}^{J} n_j = N, \quad \sum_{j=1}^{J} W_j = \sum_{i=1}^{N} i = \frac{1}{2}N(N+1).$$

当 J 个子样的容量都相等, 即 $n_j = N/J(j = 1, \cdots, J)$, 则式 (12.8.53) 简化为

$$H = \frac{12J}{N^2(N+1)}\sum_{j=1}^{J} W_j^2 - 3(N+1)$$

如果在按多个总体子样观测值的数值递增次序排列成的混合顺序统计量数列中, 出现几个相同的数值 (因而排列在一起), 这几个数值形成一个 "**结**"(tie), 每个数值对应的秩都规定为它们的顺序号的平均值, 如同威尔科克森秩和检验中的情形相同. 当结数多于 2 个, 需将统计量 H 修正为

$$H_0 = \frac{H}{C}, \quad C = 1 - \frac{\sum^{t} f_t(f_t^2 - 1)}{N(N^2-1)} \tag{12.8.54}$$

其中 t 是**结**的个数, f_t 是每个**结**中数值相同子样测量值的个数. 例如, 对混合顺序统计量数列

$$0.8,\ 0.9,\ 1.1,\ \underbrace{1.2,\ 1.2,\ 1.2},\ 1.4,\ 1.5,\ \underbrace{1.7,\ 1.7},1.9,\ 2.1,\ 2.3,\ \underbrace{2.4,\ 2.4},\ 2.7$$

有三个**结**, $f_{t_1}=3, f_{t_2}=2, f_{t_3}=2$.

$$C=1-\frac{3(9-1)+2\times 2(4-1)}{16(16^2-1)}=1-0.0088=0.9912.$$

这三个**结**对应的秩分别是 5, 9.5 和 14.5.

当 H_0 为真, 对所有 j 应有 $\overline{W_j}\sim E(\overline{W_j})$, 由式 (12.8.52) 可知, H 应接近于 0; 反之, 若 H_0 不为真, 对于不同的 j, $\overline{W_j}-E(\overline{W_j})$ 的差值有很大差别, 故 $H>0$. 这样, 统计量 H 的值对于 J 个子样是否抽自同一总体相当敏感, 可以作为 J 个子样一致性的检验统计量. 而且由以上分析可知, 对于给定的显著性水平 α, 临界域应在临界值 H_α 的上侧.

为了求出临界值 H_α, 必须知道零假设 H_0 为真 (J 个子样来源于同一总体) 时统计量 H 的概率分布. 这是一个求组合方式个数的问题, 原则上不难解出; 但当子样个数 J 增大时, 计算量迅速增大, 实际上很难实现.

克鲁斯卡尔–瓦列斯给出了子样数 $J=3$, 每个子样容量 $n_j(j=1,2,3)\leqslant 5$ 的临界值 H_α 的表 (见附表 22), 其中显著性水平 α 大都在 0.01~0.10.

当 H_0 为真, 并且各子样容量 n_j 充分大, 统计量 H 渐近地服从自由度 $J-1$ 的 χ^2 分布, 因此可利用 $\chi^2(J-1)$ 的概率密度 $f(u;J-1)$ 的积分求出 K-W 检验的临界值 $H_\alpha=\chi^2_\alpha(J-1)$, 且

$$\begin{aligned}\alpha&=\int_{\chi^2_\alpha(J-1)}^{\infty}f(u;J-1)\mathrm{d}u=1-\int_0^{\chi^2_\alpha(J-1)}f(u;J-1)\mathrm{d}u\\&=1-F(H_\alpha:J-1),\end{aligned}\tag{12.8.55}$$

其中 $F(H_\alpha:J-1)$ 表示 $\chi^2(J-1)$ 的累积分布函数值. 在实际问题中, 当 $J=3$ 且所有子样容量 $\geqslant 5$ 或 $J\geqslant 4$ 且所有子样容量 $\geqslant 4$ 时, 上述 $\chi^2(J-1)$ 近似一般是可以接受的.

对于并非所有子样容量都 $\geqslant 5$, 而从附表 22 无法查到 H_α 的情形 (如 $J>3$, 或 $J=3$ 但有一个或两个子样容量 >5), 可采用下述近似方法求得一定显著性水平 α 的临界值 H_α. 当零假设 H_0 为真, 统计量 H 的期望值 E, 方差 V 和极大值 M 分别为

$$E=J-1,$$

$$V=2(J-1)-\frac{2\left[3J^2-6J+N(2J^2-6J+1)\right]}{5N(N+1)}-\frac{6}{5}\sum_{j=1}^{J}\frac{1}{n_j}\ ,$$

$$M = \frac{N^3 - \sum_{j=1}^{J} n_j^3}{N(N+1)}; \tag{12.8.56}$$

令

$$\begin{aligned} F &= \frac{H_\alpha(M-E)}{E(M-H_\alpha)}, \\ f_1 &= \frac{2E}{MV}\left[E(M-E)-V\right], \\ f_2 &= \frac{M-E}{E}f_1, \end{aligned} \tag{12.8.57}$$

$$k_p = \frac{\left(1-\dfrac{2}{9f_2}\right)F^{1/3} - \left(1-\dfrac{2}{9f_1}\right)}{\left(\dfrac{2F^{2/3}}{9f_2}+\dfrac{2}{9f_1}\right)^{1/2}}, \tag{12.8.58}$$

于是 H_0 为真时, 统计量 $H \geqslant H_\alpha$ 的概率可表示为

$$\alpha = P(H \geqslant H_\alpha | H_0) \cong \int_{k_p}^{\infty} N(0,\ 1)\mathrm{d}x, \tag{12.8.59}$$

其中 $N(0,1)$ 是标准正态分布. 换句话说, 如果将 H 的观测值 H_{obs} 代入式 (12.8.57) 中的 H_α, 并由式 (12.8.58)、式 (12.8.59) 计算出相应的 α 值 α_{obs}, 那么若取定的显著性水平 $\alpha > \alpha_{\text{obs}}$, 则可接受 H_0; 反之, 如 $\alpha < \alpha_{\text{obs}}$, 则应拒绝零假设 H_0.

在利用附表 22 或 χ^2 近似的情形下, 给定显著性水平 α, 由附表 22 或式 (12.8.55) 可求得临界值 H_α, 利用各子样观测值按式 (12.8.53) 计算统计量 H 的观测值 H_{obs}, 若有

$$H_{\text{obs}} > H_\alpha,$$

则在显著性水平 α 上拒绝零假设; 反之, 则应认为各组观测来源于同一总体.

12.8.7　多个离散总体子样的 χ^2 检验

12.8.5 节和 12.8.6 节讨论的秩检验方法仅适用于待检验一致性的两个或多个总体均为连续分布的情形. 当被比较的总体是离散分布, 各个总体的子样中可以有多个数值相同的观测值, 每个观测值对应的秩不是唯一的. 如果根据子样的秩构建统计量, 它的分布难以导出, 因而秩检验无法应用. 同样, 如果多个连续总体的子样观测值在进行相互比较之前已经归并为直方图型数据, 比较这 J 个直方图的一致性与比较 J 个离散总体子样的一致性问题相当, 也不能用秩检验, 而可应用 χ^2 检验.

设有 J 个直方图, 它们的 I 个子区间是相同的. 对于第 j 个直方图 $(j=1,\cdots,J)$, 假定观测值落在第 $i(i=1,\cdots,I)$ 个子区间内的概率用 p_{ij} 表示, 归一化条件要求

$$\sum_{i=1}^{I} p_{ij}=1, \quad j=1,\cdots,J. \tag{12.8.60}$$

各子区间内的观测频数用 $n_{ij}(i=1,\cdots,I)$ 表示, 第 j 个直方图中观测到的事例总数为

$$\sum_{i=1}^{I} n_{ij}=n_{\cdot j}, \quad j=1,\cdots,J. \tag{12.8.61}$$

对于全部 J 个直方图, 观测事例总数为

$$n=\sum_{i=1}^{J} n_{\cdot j}=\sum_{j=1}^{J}\sum_{i=1}^{I} n_{ij}. \tag{12.8.62}$$

待检验的问题是, 这 J 个直方图所代表的总体具有相同的分布, 这相当于对所有 J 个总体, 在子区间 $i(i=1,\cdots,I)$ 中事件发生的概率相同, 即

$$H_0: p_{i1}=p_{i2}=\cdots=p_{iJ}, \quad i=1,\cdots,I. \tag{12.8.63}$$

当 H_0 为真, 记 $p_{i\cdot}=p_{i1}=\cdots=p_{iJ}$, $i=1,\cdots,I$. 由极大似然法可求得它们的极大似然估计是 J 个直方图中第 i 个子区间内的平均观测频率值, 即

$$\hat{p}_{i\cdot}=\sum_{j=1}^{J} n_{ij}/n, \quad i=1,\cdots,I. \tag{12.8.64}$$

由式 (12.8.62) 的约束可知

$$\sum_{i=1}^{I}\hat{p}_{i\cdot}=\sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij}/n=1,$$

因此, $\hat{p}_{i\cdot}$ $(i=1,\cdots,I)$ 中只有 $(I-1)$ 个是独立的. 令

$$X^2\equiv\sum_{j=1}^{J}\sum_{i=1}^{I}\frac{(n_{ij}-\hat{p}_{i\cdot}n_{\cdot j})^2}{\hat{p}_{i\cdot}n_{\cdot j}}, \tag{12.8.65}$$

它是 J 个直方图共 IJ 个子区间内观测频数 n_{ij} 和 H_0 为真时的估计频数 $\hat{p}_{i\cdot}\, n_{\cdot j}$ 离差的平方, 除以估计频数并求和. 这个统计量与单子样 χ^2 检验中的检验统计量 X^2 有相似的形式. 当 H_0 为真, 对不同的直方图, 第 i 个子区间内事件发生的概率 $p_{i1},p_{i2},\cdots,p_{iJ}$ 趋向于同一估计值 $\hat{p}_{i\cdot}$, 离差 $n_{ij}-\hat{p}_{i\cdot}\, n_{\cdot j}$ 趋向于零, 因此, X^2 趋近

于零; 反之, 离差则增大. 因此, X^2 对 J 个直方图所服从的总体之间的一致性敏感, 可作为检验它们的一致性的统计量. 而且由以上讨论可以确定, 临界域应在临界值的上侧.

若 H_0 为真, 且所有直方图的各子区间内事例数的估计值 $\hat{p}_{i\cdot}\ n_{\cdot j}$ 足够大, 式 (12.8.65) 定义的统计量 X^2 近似地满足χ^2 分布, 自由度为 $(I-1)(J-1)$, 因为总共 IJ 个观测值 n_{ij} 中存在 J 个约束条件 (见式 (12.8.61)), 故只有 $IJ-J$ 个独立观测, 并且有 $(I-1)$ 个独立的待估计参数 (见式 (12.8.64)), 自由度数等于独立观测数减去独立参数的个数, 于是

$$\nu = IJ - J - (I-1) = (I-1)(J-1).$$

对于给定的显著性水平 α, 由 $\chi^2(\nu)$ 分布查得对应的临界值 $\chi^2_\alpha(\nu)$

$$1-\alpha = F(\chi^2_\alpha(\nu);\nu),$$

其中 $F(u;\nu)$ 是 $\chi^2(\nu)$ 分布的累积分布函数. 根据观测数据 $n_{ij}(i=1,\cdots,I;\ j=1,\cdots,J)$ 和式 (12.8.65) 可算出 X^2 的观测值 $X^2_{\rm obs}$, 如

$$X^2_{\rm obs} > \chi^2_\alpha(\nu),$$

则在显著性水平 α 上拒绝 J 个直方图一致性的零假设.

在 $X^2_{\rm obs} > \chi^2_\alpha(\nu)$ 的情形下, J 个直方图的一致性被排除, 如果检查各个直方图数据对 $X^2_{\rm obs}$ 的贡献, 有时可能发现某个直方图的贡献异乎寻常地大. 将该直方图的数据摒弃后, 重新作 χ^2 检验, 可能其余 $J-1$ 个直方图在给定显著性 α 上是一致的. 这时应对贡献过大的那个直方图数据做仔细的检查, 从物理上找出数据可能错误的原因并加以修正.

对于 J 个离散总体子样的一致性检验问题, 如果 J 个总体含有相同的 I 个元素, 第 $j(j=1,\cdots,J)$ 个总体第 $i(i=1,\cdots,I)$ 个元素的观测频数记为 n_{ij}, 则可以比照直方图一致性检验的步骤进行.

例 12.20　*多组实验数据之间的一致性*

某两种粒子碰撞后有五种不同的末态, 即可看成五种不同的反应类型. 四个实验组分别对这粒子碰撞做实验研究, 观测到的不同反应类型的事例数见表 12.17, 问他们的实验数据是否互相一致.

这是四个离散分布的一致性检验问题. 将上表中 n_{ij} 的数据代入式 (12.8.61), 式 (12.8.64), 式 (12.8.65), 求得 $X^2_{\rm obs}=8.64$, 自由度 $\nu=(5-1)(4-1)=12$. 由附表 7 查得, $\alpha=0.05$ 和 0.10 时临界值分别为

$$\chi^2_\alpha = 21.03 \text{ 和 } 18.55.$$

表 12.17 四个实验组数据的一致性检验

反应类型 \ 实验组	1	2	3	4	5
A	202	150	107	51	18
B	152	131	70	48	17
C	189	161	108	42	25
D	105	78	52	32	12

显然, 在通常选取的显著性水平 $\alpha = 0.05 \sim 0.10$ 上不能排除四个实验组数据相互一致的原假设.

第 13 章　贝叶斯统计

贝叶斯统计得名于英国统计学家 Thomas Bayes, 从其 1763 年发表的论文《论机遇问题的求解》[78] 开始, 逐步发展成为一种统计推断的体系. 近几十年来, 统计学中的贝叶斯 (Bayes) 学派有了重大发展, 如今贝叶斯学派已成为与经典学派 (即频率学派) 并列的两大学派之一.

前面几章 (7~12 章) 介绍的是经典统计学的基本内容. 经典统计推断中只利用两类信息：总体信息 (总体服从的分布) 和样本信息 (即样本观测值 $X = (X_1, \cdots, X_n)^T$). 在经典统计中待估计的参数 ϑ 被认为是一未知常数, 在抽样之前对其一无所知, 关于参数 ϑ 的信息和知识完全来源于总体信息和样本信息. 但是在许多实际问题中, 可能在抽样之前就有了一些关于参数 ϑ 的信息, 它主要来自于某些经验知识或历史资料. 例如工厂一般每天都要对其主要产品抽检其废品率, 经过一段时间就积累了对过去产品废品率的一定的知识. 对当今的产品废品率这一参数进行估计时, 若能利用这些信息无疑是有帮助的. 由于这些信息发生在对当今废品率进行估计时的随机抽样之前, 故称为先验信息. 利用总体信息、样本信息和待估计参数的先验信息进行统计推断就是贝叶斯统计的基本特征.

与经典统计中待估计参数 ϑ 被认为是一未知常数不同, 贝叶斯统计认为任一未知参数都具有不确定性, 故应该视为一随机变量. 例如工厂测量某一产品当天的废品率 p, 每天都有一个 p 值, 其值因为某些随机因素的影响而逐日有所波动, 用一个分布来描述这种波动是合理的, 因此废品率应当考虑为一个随机变量.

经典统计在各个领域获得了广泛的应用, 但同时也存在一些不足, 例如估计问题中的枢轴变量和假设检验问题中的检验统计量的寻找和分布函数的确定比较困难, 置信区间的含义模糊, 等等. 而在这些方面, 贝叶斯统计却不存在这类困难. 经典统计所涉及的主要应用如参数估计和假设检验, 贝叶斯统计均可从自身的理论出发导出相应的结论. 在一些应用领域如社会科学、经济商业活动、决策问题中, 贝叶斯统计比经典统计更具优势, 因而得到了广泛的应用.

本章主要介绍贝叶斯统计的基本理论方法及其在参数估计和假设检验中的应用. 关于贝叶斯统计的基本理论方法可参阅 Jeffreys[79], Savage[80], de Finetti[81] 等的著作和 Cox[82] 的论文, 贝叶斯方法的应用则可参阅 Lindley[83], Lee[84] 和 Sivia[85] 的著作.

13.1 频率概率和贝叶斯概率

如 1.2 节所述, 经典学派对于事件 A 的概率定义为随机试验次数 N 趋向无穷大的极限情形下事件 A 出现的频率：

$$P(A) = \lim_{N\to\infty} \frac{n}{N},$$

式中 n 是 N 次随机试验中事件 A 出现的次数. 这一定义的优点是概率具有客观性, 不依赖于观测主体；并且具有精确性, 原则上只要 N 足够大就可以达到任意精度. 但它亦有局限性, 只适用于在相同条件下试验可以多次重复的随机事件.

事件 A 的**贝叶斯概率**则定义为观测主体对一个单次试验中事件 A 发生可能性大小的信任程度, 或称**信度**(degree of belief), 它表示观测主体对于事件发生的不确定性的一种主观判断. 在贝叶斯统计中, 样本空间的元素 (基本事件) 相应于一组互斥的基本假设或陈述 (hypothesis or proposition), 并且在一次随机试验中必有且只有一种基本假设为真, 即这组基本假设是完备的. 一个假设可以包含一个或一个以上的基本假设. 与假设 (事件)A 相关的贝叶斯概率定义为

$$P(A) = (\text{观测者})\text{相信假设}A\text{为真的信度}.$$

从定义知道, 贝叶斯概率带有观测主体的主观认知成分, 因而是一种**主观概率**. 应当指出, 与经典概率一样, 贝叶斯概率也需服从 1.2 节中阐明的概率的公理化定义的要求.

贝叶斯概率的定义适用的随机现象更为广泛, 不但适用于可重复的现象 (例如掷骰子试验), 也可以应用于不可重复的现象 (比如, 明天将会下雨这一事件的概率). 在掷骰子试验中, {出现 1 点, 出现 2 点, $\cdots$, 出现 6 点}构成随机试验的完备组, 但骰子出现 i 点 $(i = 1, \cdots, 6)$ 的贝叶斯概率 $P(i)$ 并不需要进行大量随机试验来确定, 而可以按照观测者的主观经验来确定.

贝叶斯概率的定义是由于客观存在和客观需要.

例如, 在高考前, 一位教师认为学生甲考取大学本科的概率是 0.9, 学生乙考取大学本科的概率是 0.6, 这是教师根据自己对学生甲、乙的了解和对高考的了解分别给出的个人信念. 显然实际结果是学生或者考上了, 或者考不上, 两者只能取其一. 又比如, 甲、乙、丙 3 人早上一起赴某地办事, 晚上 18 时回来, 为了决定是否带伞, 他们根据自己的知识和经验对于

$$A = \text{今天18时前会下雨}$$

这一事件会发生的可能性做了估计, 分别为 0.2, 0.3, 0.85, 结果甲、乙决定带伞, 而丙则不带. 这些例子经常遇到. 他们给出的主观概率决不是随意的, 而是基于当事

人对所考察的事件的了解和经验, 并对历史和新的信息进行分析的基础上得到的判断. 这样的主观概率有其一定的客观背景, 因此, 应当把主观概率与主观臆造区别开来.

贝叶斯概率的局限性在于带有观测主体的成分, 它不是在坚实的客观基础上为人们所公认, 因而可能接近、也可能偏离客观实际, 根据贝叶斯推断所做出的决策有可能不准确或不正确, 也就是存在决策错误的风险. 例如上例中不带伞的丙冒有被雨淋的风险, 如果天确实下雨而且他恰巧行走在路上而挨了雨淋, 那他的决策就是错误的. 因此贝叶斯统计推断中决策的风险评估以及对于风险发生时的可忍受程度的评估极其重要.

许多随机现象不可能重复, 特别是在经济领域和决策问题中, 因而无法用频率法去确定事件概率. 从这一意义上看, 主观概率至少可在频率法不适用的情形下来使用概率的概念, 并运用其统计推断方法.

13.2 贝叶斯公式和贝叶斯统计模型

13.2.1 贝叶斯公式

从概率论中的贝叶斯公式谈起. 设事件 $B_1,\cdots,B_k$ 是随机试验样本空间 Ω 的互不相容事件的完备组, A 是 Ω 的任一事件, 则事件 A 发生条件下事件 B_i 发生的条件概率用贝叶斯公式表示：

$$P(B_i|A)=\frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{k}P(A|B_j)P(B_j)},\quad i=1,\cdots,k. \tag{13.2.1}$$

在贝叶斯统计中, $P(B_j),j=1,\cdots,k$ 这一概率分布表示的是随机试验前人们对于发生事件 B_j 的概率的认识, 它反映的一般是以往知识的总结, 称为**先验概率**. 条件概率 $P(A|B_i)$ 表示新的随机试验的结果. 条件概率 $P(B_i|A)$ 称为**后验概率**, 由于事件 A 的发生, 可以对事件 B_i 的发生概率提供新的信息, 以后验概率体现出来, 是对先验概率知识的更新.

为了引入贝叶斯统计模型, 需要将式 (13.2.1) 用随机变量的形式来表示. 引入随机变量 ϑ, 它的取值为 $\vartheta_1,\cdots,\vartheta_k$, 其中 $\vartheta_j=\vartheta(B_j)$, 即事件 B_j 发生时 ϑ 取值 ϑ_j. 随机变量 ϑ 具有先验分布 $\pi(\vartheta)$：

$$\pi(\vartheta_j)=P(\vartheta=\vartheta_j)=P(B_j),\quad j=1,\cdots,k.$$

又事件 A 为一随机事件, 可定义一个随机变量 X, 使得

$$X(A)=X.$$

这样, 式 (13.2.1) 中的 $P(A|B_j)$ 可表示为

$$P(X|\vartheta_j)=P(X|\vartheta=\vartheta_j),\quad j=1,\cdots,k.$$

它代表一种样本分布, 这里样本容量 $n=1$. 这种样本分布依赖于参数 ϑ, 而且 ϑ 是随机变量. 样本分布记为 $P(X|\vartheta)$. 于是贝叶斯公式 (13.2.1) 可表示为

$$\begin{aligned}P(\vartheta_i|X)&=P(\vartheta=\vartheta_i|X)\\&=\frac{P(X|\vartheta_i)\pi(\vartheta_i)}{\sum_{j=1}^{k}P(X|\vartheta_j)\pi(\vartheta_j)},\quad i=1,\cdots,k.\end{aligned}\tag{13.2.2}$$

对于参数 ϑ 和随机变量 X 的不同类型, 贝叶斯公式有不同的形式.

(1) ϑ 和 X 皆为离散型

这时参数 ϑ 的先验分布 $\pi(\vartheta)$ 的概率分布律为

$$\begin{aligned}&\vartheta:\quad \vartheta_1\cdots\vartheta_j\cdots,\\&\pi(\vartheta):\pi_1\cdots\pi_j\cdots.\end{aligned}\tag{13.2.3}$$

X 的容量 n 的样本 $\boldsymbol{X}=(X_1,\cdots,X_n)^T$ 亦是离散型, 当 $\vartheta=\vartheta_j$ 时, 其概率分布是 $P(\boldsymbol{X}|\vartheta_j)$, 则 ϑ 的后验分布是

$$P(\vartheta_i|\boldsymbol{X})=\frac{P(\boldsymbol{X}|\vartheta_i)\pi_i}{\sum_{j}P(\boldsymbol{X}|\vartheta_j)\pi_j},\quad i=1,2,\cdots.\tag{13.2.4}$$

(2) ϑ 和 X 皆为连续型

设参数 ϑ 具有先验概率密度 $\pi(\vartheta)$, 在给定 ϑ 时样本 X 的条件概率密度是 $f(\boldsymbol{X}|\vartheta)$. 这时 $(\vartheta,\boldsymbol{X})$ 的联合概率密度是

$$f(\boldsymbol{X}|\vartheta)\pi(\vartheta).$$

而 $\boldsymbol{X}$ 的边沿概率密度为

$$q(\boldsymbol{X})=\int_{\Theta}f(\boldsymbol{X}|\vartheta)\pi(\vartheta)\mathrm{d}\vartheta,$$

其积分区域 Θ 是参数空间. 边沿概率密度是样本条件概率密度按参数先验分布求平均, 与 ϑ 无关.

有了 $(\boldsymbol{X},\vartheta)$ 的联合概率密度和 $\boldsymbol{X}$ 的边沿密度, 容易求得给定 $\boldsymbol{X}$ 条件下 ϑ 的条件分布：

$$h(\vartheta|\boldsymbol{X})=\frac{f(\boldsymbol{X}|\vartheta)\pi(\vartheta)}{q(\boldsymbol{X})}=\frac{f(\boldsymbol{X}|\vartheta)\pi(\vartheta)}{\int_{\Theta}f(\boldsymbol{X}|\vartheta)\pi(\vartheta)\mathrm{d}\vartheta}.\tag{13.2.5}$$

$h(\vartheta|\boldsymbol{X})$ 称为参数 ϑ 的**后验概率密度**.

(3) ϑ 为离散型, X 为连续型

这时参数 ϑ 具有先验概率分布律式 (13.2.3), 给定 ϑ 时样本 $\boldsymbol{X}$ 的条件概率密度是 $f(\boldsymbol{X}|\vartheta)$. 当给定 $\boldsymbol{X}$ 条件下 ϑ_i 的条件概率是

$$P\left(\vartheta_i|\boldsymbol{X}\right)=\frac{f\left(\boldsymbol{X}|\vartheta_i\right)\pi\left(\vartheta_i\right)}{\sum_j f\left(\boldsymbol{X}|\vartheta_j\right)\pi\left(\vartheta_j\right)},\quad i=1,2,\cdots. \tag{13.2.6}$$

(4) ϑ 为连续型, X 为离散型

这时参数 ϑ 具有先验概率密度 $\pi(\vartheta)$, 在给定 ϑ 时样本 $\boldsymbol{X}$ 的条件概率是 $P(\boldsymbol{X}|\vartheta)$. 于是给定 $\boldsymbol{X}$ 条件下 ϑ 的条件概率密度为

$$h(\vartheta|\boldsymbol{X})=\frac{P(\boldsymbol{X}|\vartheta)\pi(\vartheta)}{\int_\Theta P(X|\vartheta)\pi(\vartheta)\mathrm{d}\vartheta}. \tag{13.2.7}$$

贝叶斯公式 (13.2.4) 至式 (13.2.7) 可以统一起来. 给定 ϑ, 样本 $\boldsymbol{X}$ 是离散型时的条件概率是 $P(\boldsymbol{X}|\vartheta)$, $\boldsymbol{X}$ 是连续型时的条件概率密度是 $f(\boldsymbol{X}|\vartheta)$, 现在两种情形下都用 $f(\boldsymbol{X}|\vartheta)$ 表示, 统称为**条件分布**. 类似地, 给定 $\boldsymbol{X}$, 参数 ϑ 是离散型时的后验分布是 $P(\vartheta|\boldsymbol{X})$, ϑ 是连续型时的后验概率密度是 $h(\vartheta|\boldsymbol{X})$, 现在两种情形下都用 $h(\vartheta|\boldsymbol{X})$ 表示, 统称为**后验分布**. 先验分布的离散概率分布律 $\pi(\vartheta_i)$ 和概率密度 $\pi(\vartheta)$ 都用 $\pi(\vartheta)$ 表示, 统称为**先验分布**. 式 (13.2.5) 中的积分当 ϑ 是离散型时变为相应的求和. 于是贝叶斯公式统一为式 (13.2.5) 的形式.

13.2.2　贝叶斯统计模型和贝叶斯推断原则

基于前面的讨论, **贝叶斯统计模型**可表述为:

(1) 参数 ϑ 的参数空间 Θ 上的一个概率分布 (连续或离散)$\{\pi(\vartheta):\vartheta\in\Theta\}$ 称为 ϑ 的先验分布.

(2) 样本 $\boldsymbol{X}=(X_1,\cdots,X_n)^T$ 的条件分布族 (连续或离散)$\{f(\boldsymbol{X}|\vartheta):\vartheta\in\Theta\}$ 称为样本分布族.

(3) 先验分布 $\{\pi(\vartheta):\vartheta\in\Theta\}$ 和样本分布族 $\{f(\boldsymbol{X}|\vartheta):\vartheta\in\Theta\}$ 构成贝叶斯统计模型.

在样本观测值 $\boldsymbol{X}$ 已知的条件下, ϑ 的条件分布称为 ϑ 的后验分布. 后验分布用后验分布族 $\{h(\vartheta|\boldsymbol{X}):\vartheta\in\Theta\}$ 描述.

贝叶斯公式 (13.2.4) 到 (13.2.7) 是贝叶斯统计的基本公式. 后验分布的意义在于综合了 ϑ 的先验信息 (反映在先验分布 $\pi(\vartheta)$ 中) 和样本 $\boldsymbol{X}$ 关于 ϑ 的信息 (反映在样本分布 $f(\boldsymbol{X}|\vartheta)$ 中), 先验分布概括了随机试验前对 ϑ 的知识, 而得到样本观

测值 $\boldsymbol{X}$ 后, 对 ϑ 的认识得到了深化和更新, 集中地反映在后验分布中. 贝叶斯公式反映了先验分布到后验分布的转化. 后验分布是贝叶斯统计推断的基础, 由此引出**贝叶斯统计推断原则**：对参数 ϑ 所做的任何推断 (参数估计、假设检验等), 必须而且只能基于参数 ϑ 的后验分布族 $\{h(\vartheta|\boldsymbol{X}):\vartheta\in\Theta\}$.

与贝叶斯统计模型中参数 ϑ 视为随机变量不同, 经典的统计模型中参数 ϑ 视为未知常数或未知常数向量, 样本分布族 $\{f(\boldsymbol{X};\vartheta):\vartheta\in\Theta\}$ 是带参数 ϑ 的分布族, 不能理解为条件分布族. 经典统计的基本问题是由样本分布族 $\{f(\boldsymbol{X};\vartheta):\vartheta\in\Theta\}$ 出发, 对 ϑ 进行统计推断 (参数估计、假设检验等).

现在我们引入一个有用的概念：**密度函数的核**. 若随机变量 X 的密度函数 $f(X)=cg(X)$, 其中 c 为与 X 无关的常数, 则可记 $f(X)\propto g(X)$, $g(X)$ 称为密度函数 $f(X)$ 的核. 几个常见分布关于其参数的核如表 13.1 所示.

表 13.1　常见分布关于其参数的核

分布	概率密度	核 (关于参数 ϑ)
二项分布 $b(X;n,\theta)$	$\binom{n}{X}\theta^X(1-\theta)^{n-X}, X=0,1,\cdots,n$	$\theta^X(1-\theta)^{n-X}$
泊松分布	$\frac{1}{X!}\theta^X\mathrm{e}^{-\theta},\quad X=0,1,\cdots$	$\theta^X\mathrm{e}^{-\theta}$
指数分布	$\frac{1}{\theta}\mathrm{e}^{-\frac{X}{\theta}},\quad X\in[0,\infty)$	$\frac{1}{\theta}\mathrm{e}^{-\frac{X}{\theta}}$
正态分布 $N(X;\theta,\sigma)$	$\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(X-\theta)^2}{2\sigma^2}},\quad X\in(-\infty,\infty)$	$\mathrm{e}^{-\frac{(X-\theta)^2}{2\sigma^2}}$
正态分布 $N(X;\mu,\theta^2)$	$\frac{1}{\sqrt{2\pi}\theta}\mathrm{e}^{-\frac{(X-\mu)^2}{2\theta^2}},\quad X\in(-\infty,\infty)$	$\frac{1}{\theta}\mathrm{e}^{-\frac{(X-\mu)^2}{2\theta^2}}$
正态分布 $N(X;\mu,\theta)$	$\frac{1}{\sqrt{2\pi\theta}}\mathrm{e}^{-\frac{(X-\mu)^2}{2\theta}},\quad X\in(-\infty,\infty)$	$\frac{1}{\sqrt{\theta}}\mathrm{e}^{-\frac{(X-\mu)^2}{2\theta}}$
正态分布 $N(X;\theta_1,\theta_2^2)$	$\frac{1}{\sqrt{2\pi}\theta_2}\mathrm{e}^{-\frac{(X-\theta_1)^2}{2\theta_2^2}},\quad X\in(-\infty,\infty)$	$\frac{1}{\theta_2}\mathrm{e}^{-\frac{(X-\theta_1)^2}{2\theta_2^2}}$
正态分布 $N(X;\theta_1,\theta_2)$	$\frac{1}{\sqrt{2\pi\theta_2}}\mathrm{e}^{-\frac{(X-\theta_1)^2}{2\theta_2}},\quad X\in(-\infty,\infty)$	$\frac{1}{\sqrt{\theta_2}}\mathrm{e}^{-\frac{(X-\theta_1)^2}{2\theta_2}}$
伽马分布 $\Gamma(\theta;a,b)$	$\frac{b^a}{\Gamma(a)}\theta^{a-1}\mathrm{e}^{-b\theta},\quad \theta\in[0,\infty)$	$\theta^{a-1}\mathrm{e}^{-b\theta}$
贝塔分布 $\beta(\theta;a,b)$	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{\theta^{a-1}}{(1-\theta)^{1-b}},\quad \theta\in[0,1]$	$\theta^{a-1}(1-\theta)^{b-1}$

由式 (13.2.5) 知, 由于 $q(\boldsymbol{X})$ 与 ϑ 无关, 在计算 ϑ 的后验分布时仅起到一个归一化因子的作用, 因此在对 ϑ 进行统计推断时, 可将 $q(\boldsymbol{X})$ 省略, 也就是只需考虑后验分布密度 $h(\vartheta|\boldsymbol{X})$ 关于参数 ϑ 的核, 于是贝叶斯公式可改写成如下等价的形式：

$$h(\vartheta|\boldsymbol{X})\propto f(\boldsymbol{X}|\vartheta)\pi(\vartheta),\quad \vartheta\in\Theta. \tag{13.2.8}$$

给定样本 $\boldsymbol{X}=(X_1,\cdots,X_n)^T$, 则样本分布 $f(\boldsymbol{X}|\vartheta)$ 即是似然函数 $L(\boldsymbol{X}|\vartheta)$, 故有

$$h(\vartheta|\boldsymbol{X}) \propto L(\boldsymbol{X}|\vartheta)\pi(\vartheta), \quad \vartheta \in \Theta. \tag{13.2.9}$$

贝叶斯统计中许多情形下 $h(\vartheta|\boldsymbol{X})$ 采用这样的形式计算起来十分方便.

到目前为止, 先验分布 $\pi(\vartheta)$ 必须满足一般随机变量概率密度的要求, 即

$$\pi(\vartheta) \geqslant 0, \quad \int_\Theta \pi(\vartheta)\mathrm{d}\vartheta = 1. \tag{13.2.10}$$

但有些情况下, 后一积分值不等于 1, 甚至为无穷, 但边沿概率密度 $q(\boldsymbol{X})$ 仍为有限值, 这时式 (13.2.5) 定义的 $h(\vartheta|\boldsymbol{X})$ 作为 ϑ 的函数仍然满足概率密度的条件, 也就是说, 这样的函数 $\pi(\vartheta)$ 仍然可以作为先验分布. 由于这时 $\int_\Theta \pi(\vartheta)\mathrm{d}\vartheta \neq 1$, 它已失去了通常的概率密度的意义, 这样的 $\pi(\vartheta)$ 被称为**广义先验分布**. 用数学公式来描述, 即广义先验分布 $\pi(\vartheta)$ 满足:

$$\int_\Theta \pi(\vartheta)\mathrm{d}\vartheta = +\infty, \quad \int_\Theta f(\boldsymbol{X}|\vartheta)\,\pi(\vartheta)\mathrm{d}\vartheta < +\infty. \tag{13.2.11}$$

当利用了广义先验分布 $\pi(\vartheta)$, 当然会导致相应的后验密度不是概率密度函数了. 这样的考虑当然是有道理的. 目前的研究限定于以式 (13.2.5) 求得的后验分布 $h(\theta|\boldsymbol{X})$ 仍然为概率密度函数的广义先验分布.

13.2.3 先验分布和后验分布, 先验分布的选择

由于对参数 ϑ 所做的任何推断必须而且只能基于参数 ϑ 的后验分布 $h(\vartheta|\boldsymbol{X})$, 而它又直接取决于先验分布 $\pi(\vartheta)$, 显然先验分布的选择就成为贝叶斯统计中一个极为重要的问题. 本节讨论先验分布选择的几种重要方法和原则.

由后验分布 $h(\vartheta|\boldsymbol{X})$ 的定义式 (13.2.5) 可知, 先验分布 $\pi(\vartheta)$ 中的常数乘因子在求 $h(\vartheta|\boldsymbol{X})$ 过程中是相消的, 因此在选择先验分布 $\pi(\vartheta)$ 时可以忽略常数乘因子, 只需要考虑它关于参数 ϑ 的核.

(1) **先验分布的贝叶斯假设**

贝叶斯本人对先验分布做了如下假设: 先验分布是 “**无信息先验分布**”, 应在参数 ϑ 的取值区域内均匀分布, 即

$$\pi(\vartheta) = c \text{ 或 } \quad \pi(\vartheta) \propto 1, \quad \vartheta \in \Theta. \tag{13.2.12}$$

贝叶斯假设又可称为 “同等无知假设”, 即认为参数 ϑ 在其取值区域内取任意值有同等的可能性. 这样的先验分布除了参数的取值区域之外没有给出 ϑ 的信息, 因而称为无信息先验分布. 显然, 同等无知假设是在对参数确实没有什么信息时不得已采用的方法.

当 ϑ 的取值区域为 $(-\infty,+\infty)$ 时, 无信息先验分布 $\pi(\vartheta)$ 为广义先验分布, 这时后验分布的核按式 (13.2.9) 为

$$h(\vartheta|\boldsymbol{X}) \propto L(\boldsymbol{X}|\vartheta)\pi(\vartheta) = L(\boldsymbol{X}|\vartheta). \tag{13.2.13}$$

由于似然函数是归一化的, 故似然函数可视为贝叶斯假设下的后验分布形式.

例 13.1 用贝叶斯假设求正态分布均值的后验分布

设 $\boldsymbol{X}=(X_1,\cdots,X_n)^T$ 为正态总体 $N(\mu,\sigma^2)$ 的随机子样, σ^2 已知, μ 为未知参数. 当利用先验分布的贝叶斯假设, 后验分布正比于似然函数 $L(\boldsymbol{X}|\mu)$:

$$\begin{aligned}L(\boldsymbol{X}|\mu) &= \prod_{i=1}^{n} N(X_i;\mu,\sigma^2) \propto \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\mu-X_i)^2\right\} \\ &\propto \exp\left\{-\frac{n}{2\sigma^2}(\mu^2-2\mu\bar{X}+\bar{X}^2)\right\}\cdot\exp\left\{-\frac{n}{2\sigma^2}\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2-\bar{X}^2\right)\right\}.\end{aligned}$$

式中 $\bar{X}=\sum\limits_{i=1}^{n}X_i\Big/n$ 为子样平均. 略去与未知参数 μ 无关的系数而只考虑似然函数的核, 有

$$L(\boldsymbol{X}|\mu) \propto \exp\left\{-\frac{1}{2\sigma^2/n}(\bar{X}-\mu)^2\right\}.$$

可以看到, 该函数是均值 μ、方差 σ^2/n 的正态分布. 我们知道子样平均 $\bar{X}$ 服从均值 μ、方差 σ^2/n 的正态分布, 所以 $\boldsymbol{X}$ 的联合概率密度正比于子样平均 $\bar{X}$ 的概率密度, 即

$$f(\bar{X}|\mu) \propto L(\boldsymbol{X}|\mu) \propto \exp\left\{-\frac{1}{2\sigma^2/n}(\bar{X}-\mu)^2\right\},$$

它就是所要求的后验分布 $h(\mu|\boldsymbol{X})$, 或写为 $h(\mu|\bar{X})$:

$$h(\mu|\bar{X}) = N\left(\bar{X},\frac{\sigma^2}{n}\right).$$

(2) **不变先验分布**

贝叶斯假设是在对参数 ϑ 无信息条件下取均匀分布为先验分布. 事实上, 无信息分布的选取与参数在总体密度函数中的地位有关联, 对于总体密度函数中地位不同的参数一概取均匀分布作为先验分布不一定是恰当的做法. 所谓的**不变先验分布**就是要求先验分布对于参数在某种变换下具有不变性, 具有这种不变性的先验分布可能更为合理. 在数学上, 就是在群的作用下具有不变性.

我们首先来介绍函数的位置参数和尺度参数的概念.

设总体 X 的概率密度函数为 $f(X|\vartheta)$, 参数为 ϑ. 若 $f(X|\vartheta)$ 可表示为

$$f(X|\vartheta) \propto g(X-\vartheta), \quad \vartheta \in (-\infty, \infty), \tag{13.2.14}$$

g 是任意函数, 即 $f(X|\vartheta)$ 仅是 $X-\vartheta$ 的函数, 则 ϑ 称为**位置参数**.

若

$$f(X|\vartheta) \propto \frac{1}{\vartheta} g\left(\frac{X}{\vartheta}\right), \quad \vartheta > 0, \tag{13.2.15}$$

则 ϑ 称为**尺度参数**.

若 $\vartheta = (\vartheta_1, \vartheta_2)^T$, 且

$$f(X|\vartheta) \propto \frac{1}{\vartheta_2} g\left(\frac{X-\vartheta_1}{\vartheta_2}\right), \quad \vartheta_1 \in (-\infty, +\infty), \vartheta_2 > 0, \tag{13.2.16}$$

则 ϑ 称为**位置－尺度参数**.

位置参数和尺度参数的不变先验分布可推导如下.

(a) 对于位置参数$\boldsymbol{\vartheta}$

设总体 X 的概率密度为式 (13.2.14) 所示的 $f(X|\vartheta)$. 令随机变量 $Y = X + c$, 据随机变量概率密度与随机变量函数概率密度的关系式：

$$g(y) = f(x(y))|\frac{\mathrm{d}x(y)}{\mathrm{d}y}|.$$

Y 的概率密度为 $g(Y-c-\vartheta)$, 与总体 X 的概率密度 $f(X|\vartheta) \propto g(X-\vartheta)$ 相比较, 这相当于将参数 ϑ 平移到 $\vartheta+c$. 显然, 用总体 X 的子样对参数 ϑ 做估计与用总体 Y 的子样对参数 $\vartheta+c$ 做估计是相同的, 即 ϑ 和 $\vartheta+c$ 的先验概率密度应该是一样的. 因此, 其先验分布应在参数位置平移变换下保持不变, 即要求 $\pi(\vartheta) = \pi(\vartheta+c)$, c 为任意常数. 满足这一要求的 $\pi(\vartheta)$ 只能是常数, 即

$$\pi(\vartheta) \propto 1. \tag{13.2.17}$$

这表明, 对于位置参数, 其先验分布应服从贝叶斯假设.

(b) 对于尺度参数$\boldsymbol{\vartheta}$

设总体 X 的概率密度为式 (13.2.15) 所示的 $f(X|\vartheta) \propto \dfrac{1}{\vartheta} g\left(\dfrac{X}{\vartheta}\right)$. 令随机变量 $Y = cX$, Y 的概率密度为 $\dfrac{1}{c\vartheta} g\left(\dfrac{Y}{c\vartheta}\right)$, 与总体 X 的概率密度相比较, 这相当于将参数 ϑ 变换为 $c\vartheta$. 显然, 用总体 X 的子样对参数 ϑ 做估计与用总体 Y 的子样对参数 $c\vartheta$ 做估计是相同的, 即 ϑ 和 $c\vartheta$ 的先验概率密度应该是一样的. 因此, 其先验分布应在参数尺度变换下保持不变, 即要求 $\pi(\vartheta) = \pi(c\vartheta)$, $c > 0$. 满足这一要求的 $\pi(\vartheta)$ 为

$$\pi(\vartheta) \propto 1/\vartheta. \tag{13.2.18}$$

(c) 对于位置一尺度参数$\vartheta=(\vartheta_1,\vartheta_2)^T$,

可证明其先验分布应为[86]：

$$\pi(\vartheta)=\pi(\vartheta_1)\cdot\pi(\vartheta_2)\propto 1\cdot\frac{1}{\vartheta_2}=\frac{1}{\vartheta_2}, \tag{13.2.19}$$

即 ϑ_1、ϑ_2 相互独立, 且 $\pi(\vartheta_1)\propto 1$, $\pi(\vartheta_2)\propto 1/\vartheta_2$.

例 13.2 正态分布参数的不变先验分布

正态总体的概率密度函数为

$$f(X|\vartheta)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2\right],$$

若 σ 为常数, μ 是其唯一 (待定) 参数, 由式 (13.2.14) 知 μ 是位置参数, 不变性要求立即得到 μ 的先验分布为 $\pi(\mu)\propto 1$.

若 μ 为常数, σ 是其唯一 (待定) 参数, 由式 (13.2.15) 知 σ 是尺度参数, 不变性要求立即得到 σ 的先验分布为 $\pi(\sigma)\propto 1/\sigma$.

当 σ、μ 皆是待定参数, 由式 (13.2.16) 知 $\vartheta=(\mu,\sigma)^T$ 是位置一尺度参数, 其不变先验分布为 $\pi(\mu)\propto 1$, $\pi(\sigma)\propto 1/\sigma$.

例 13.3 指数分布参数的不变先验分布

指数总体的概率密度函数为

$$f(X|\mu)=\frac{1}{\mu}\mathrm{e}^{-X/\mu},\quad \mu>0.$$

由式 (13.2.9) 知 μ 是尺度参数, 不变性要求立即得到 μ 的先验分布为 $\pi(\mu)\propto 1/\mu$.

(3) **Jeffreys 原则**

先验分布的贝叶斯假设中存在一个矛盾, 即若先验分布对 ϑ 是同等无知的, 那么它对 $\vartheta^2,\vartheta^3,\cdots$ 或更一般地对 ϑ 的任意函数 $g(\vartheta)$ 也应当是同等无知的. 如果对参数 ϑ 选用均匀分布, 则当 ϑ 的函数 $g(\vartheta)$ 作为参数时, 也应选用均匀分布. 然而, 当 ϑ 为均匀分布时, 一般地 $g(\vartheta)$ 不为均匀分布；反之亦然. 克服这一矛盾的方法之一是 Jeffreys[79] 提出的构造无信息先验分布的原则, 它被广泛地应用于先验分布的选取.

Jeffreys 原则可陈述如下：若选取 ϑ 的先验分布为 $\pi(\vartheta)$, 则按 Jeffreys 原则选定 $\eta=g(\vartheta)$ 的先验分布应是 $\pi_g(\eta)$, 它满足如下关系

$$\pi(\vartheta)=\pi_g(g(\vartheta))\left|\frac{\partial g(\vartheta)}{\partial\vartheta}\right|; \tag{13.2.20}$$

对于参数为多维变量的情形 $\boldsymbol{\vartheta}=(\vartheta_1,\cdots,\vartheta_k)^T$, 且 $\boldsymbol{\eta}=\boldsymbol{g}(\boldsymbol{\vartheta})$ 与 $\boldsymbol{\vartheta}$ 有相同的维数 k, 则 $\boldsymbol{\vartheta}$ 的先验分布 $\pi(\boldsymbol{\vartheta})$ 与 $\boldsymbol{\eta}=\boldsymbol{g}(\boldsymbol{\vartheta})$ 的先验分布 $\pi_{\boldsymbol{g}}(\boldsymbol{\eta})$ 之间应满足如下关系

$$\pi(\boldsymbol{\vartheta})=\pi_{\boldsymbol{g}}(\boldsymbol{g}(\boldsymbol{\vartheta}))\left|\frac{\partial\boldsymbol{g}(\boldsymbol{\vartheta})}{\partial\boldsymbol{\vartheta}}\right|, \tag{13.2.21}$$

其中 $\dfrac{\partial \boldsymbol{g}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}$ 为 $k \times k$ 矩阵, 矩阵元为 $\dfrac{\partial g(\vartheta_i)}{\partial \vartheta_j}, i, j = 1, \cdots, k$, $\left|\dfrac{\partial \boldsymbol{g}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}\right|$ 表示矩阵行列式的绝对值. 当满足 Jeffreys 原则时, 用 ϑ 和用 ϑ 的函数 $g(\vartheta)$ 导出的先验分布总是一致的, 因此 Jeffreys 原则是一种不变原则.

问题在于如何找到满足式 (13.2.20) 或式 (13.2.21) 的先验分布 $\pi(\vartheta)$. Jeffreys 利用**费歇尔 (Fisher) 信息量**的一个不变性, 找到了这样的 $\pi(\vartheta)$.

对于 ϑ 为一维的情形,

$$\pi(\vartheta) \propto |I(\vartheta)|^{1/2}, \tag{13.2.22}$$

$I(\vartheta)$ 是 Fisher 信息量, 定义为

$$I(\vartheta) = E\left\{\left[\frac{\partial \ln L(\boldsymbol{X}|\vartheta)}{\partial \vartheta}\right]^2\right\} = \int \left[\frac{\partial \ln L(\boldsymbol{X}|\vartheta)}{\partial \vartheta}\right]^2 L(\boldsymbol{X}|\vartheta)\mathrm{d}\boldsymbol{X}, \tag{13.2.23}$$

式中 $\boldsymbol{X} = (X_1, \cdots, X_n)^T$ 是总体 X 的容量 n 的样本观测值.

对于 $\boldsymbol{\vartheta}$ 为 k 维矢量的情形,

$$\pi(\boldsymbol{\vartheta}) \propto |I(\boldsymbol{\vartheta})|^{1/2}, \tag{13.2.24}$$

$I(\boldsymbol{\vartheta})$ 是 $k \times k$**Fisher 信息矩阵**, 其矩阵元为

$$\begin{aligned}[I(\boldsymbol{\vartheta})]_{ij} &= E\left\{\frac{\partial \ln L(\boldsymbol{X}|\boldsymbol{\vartheta})}{\partial \vartheta_i} \cdot \frac{\partial \ln L(\boldsymbol{X}|\boldsymbol{\vartheta})}{\partial \vartheta_j}\right\} \\ &= \int \left[\frac{\partial \ln L(\boldsymbol{X}|\boldsymbol{\vartheta})}{\partial \vartheta_i} \cdot \frac{\partial \ln L(\boldsymbol{X}|\boldsymbol{\vartheta})}{\partial \vartheta_j}\right] L(\boldsymbol{X}|\boldsymbol{\vartheta})\mathrm{d}\boldsymbol{X},\end{aligned} \tag{13.2.25}$$

$|I(\boldsymbol{\vartheta})|^{1/2}$ 表示矩阵 $I(\boldsymbol{\vartheta})$ 行列式的平方根.

对于总体 X 容量 n 的样本 $\boldsymbol{X} = (X_1, \cdots, X_n)^T$ 的信息量, 由于 $X_i, i = 1, \cdots, n$ 是独立同分布的随机变量, 信息量有可相加性 [91], 即

$$I(\boldsymbol{\vartheta}) = nI_1(\boldsymbol{\vartheta}), \tag{13.2.26}$$

式中 $I(\boldsymbol{\vartheta})$ 是式 (13.2.32) 或式 (13.2.25) 定义的样本 $\boldsymbol{X} = (X_1, \cdots, X_n)^T$ 的信息量, $I_1(\boldsymbol{\vartheta})$ 定义为总体 X 的一个样本值的信息量

$$I_1(\vartheta) = E\left\{\left[\frac{\partial \ln f(X|\vartheta)}{\partial \vartheta}\right]^2\right\} = \int \left[\frac{\partial \ln f(X|\vartheta)}{\partial \vartheta}\right]^2 f(X|\vartheta)\mathrm{d}X, \tag{13.2.27}$$

或

$$[I_1(\boldsymbol{\vartheta})]_{ij} = E\left\{\frac{\partial \ln f(X|\boldsymbol{\vartheta})}{\partial \vartheta_i} \cdot \frac{\partial \ln f(X|\boldsymbol{\vartheta})}{\partial \vartheta_j}\right\}$$

$$= \int \left[\frac{\partial \ln f(X|\boldsymbol{\vartheta})}{\partial \vartheta_i} \cdot \frac{\partial \ln f(X|\boldsymbol{\vartheta})}{\partial \vartheta_j} \right] f(X|\boldsymbol{\vartheta}) \mathrm{d}X, \tag{13.2.28}$$

在满足总体 X 的取值空间与参数 $\boldsymbol{\vartheta}$ 无关, 且其样本的似然函数 $L(\boldsymbol{X}|\boldsymbol{\vartheta})$ 的导数运算 $\partial^2/\partial\vartheta_i\partial\vartheta_j$ 和在样本空间中的积分运算 $\int \mathrm{d}\boldsymbol{X}$ 可以互易的条件下, 容易证明式 (13.2.32)、式 (13.2.25) 可以改写为

$$I(\vartheta) = -E\left\{ \frac{\partial^2[\ln L(\boldsymbol{X}|\vartheta)]}{\partial \vartheta^2} \right\}, \tag{13.2.29}$$

$$[I(\boldsymbol{\vartheta})]_{ij} = -E\left\{ \frac{\partial^2[\ln L(\boldsymbol{X}|\boldsymbol{\vartheta})]}{\partial \vartheta_i \partial \vartheta_j} \right\}. \tag{13.2.30}$$

相应地, 对于总体 X 的一个样本值的信息量, 我们有

$$I_1(\vartheta) = -E\left\{ \frac{\partial^2 \ln f(X|\vartheta)}{\partial \vartheta^2} \right\} = -\int \frac{\partial^2 \ln f(X|\vartheta)}{\partial \vartheta^2} \cdot f(X|\vartheta) \mathrm{d}X, \tag{13.2.31}$$

$$[I_1(\boldsymbol{\vartheta})]_{ij} = -E\left\{ \frac{\partial^2 \ln f(X|\boldsymbol{\vartheta})}{\partial \vartheta_i \partial \vartheta_j} \right\} = -\int \frac{\partial^2 \ln f(X|\boldsymbol{\vartheta})}{\partial \vartheta_i \partial \vartheta_j} \cdot f(X|\boldsymbol{\vartheta}) \mathrm{d}X. \tag{13.2.32}$$

通常用下式确定先验分布运算较为简单:

$$\pi(\vartheta) \propto |I_1(\vartheta)|^{1/2}, \tag{13.2.33}$$

$$\pi(\boldsymbol{\vartheta}) \propto |I_1(\boldsymbol{\vartheta})|^{1/2}, \tag{13.2.34}$$

其中, $I_1(\vartheta)$ 和 $I_1(\boldsymbol{\vartheta})$ 由式 (13.2.27)、式 (13.2.28) 或式 (13.2.31)、式 (13.2.32) 确定.

下面我们用 Jeffreys 原则来确定几种总体分布的先验分布.

例 13.4 指数分布参数的 Jeffreys 先验分布.

由式 (13.2.31) 得

$$\begin{aligned} I_1(\lambda) &= -E\left\{ \frac{\partial^2 \ln f(X|\lambda)}{\partial \lambda^2} \right\} = -E\left\{ \frac{\partial^2}{\partial \lambda^2}(\ln \lambda - \lambda X) \right\} \\ &= \int_0^\infty \frac{1}{\lambda} \mathrm{e}^{-\lambda X} \mathrm{d}X = \frac{1}{\lambda^2}; \end{aligned}$$

由式 (13.2.27) 得

$$\begin{aligned} I_1(\lambda) &= E\left\{ \left[\frac{\partial \ln f(X|\lambda)}{\partial \lambda} \right]^2 \right\} = E\left\{ \left[\frac{\partial}{\partial \lambda}(\ln \lambda - \lambda X) \right]^2 \right\} \\ &= \int_0^\infty \left(\frac{1}{\lambda} - X \right)^2 \lambda \mathrm{e}^{-\lambda X} \mathrm{d}X = \frac{1}{\lambda^2} - \frac{2}{\lambda^2} + \frac{2}{\lambda^2} = \frac{1}{\lambda^2}; \end{aligned}$$

由式 (13.2.29) 得

$$I(\lambda)=-E\left\{\frac{\partial^2\ln L(\boldsymbol{X}|\lambda)}{\partial\lambda^2}\right\}=-E\left\{\frac{\partial^2}{\partial\lambda^2}\ln\left(\prod_{i=1}^n\lambda\mathrm{e}^{-\lambda X_i}\right)\right\}$$
$$=-E\left\{\frac{\partial^2}{\partial\lambda^2}\sum_{i=1}^n(\ln\lambda-\lambda X_i)\right\}=\frac{n}{\lambda^2};$$

由式 (13.2.23) 得

$$I(\lambda)=E\left\{\left[\frac{\partial\ln L(\boldsymbol{X}|\lambda)}{\partial\lambda}\right]^2\right\}=E\left\{\left[\frac{\partial}{\partial\lambda}\left(\ln\prod_{i=1}^n\lambda\mathrm{e}^{-\lambda X_i}\right)\right]^2\right\}=E\left\{\left[\sum_{i=1}^n\left(\frac{1}{\lambda}-X_i\right)\right]^2\right\}$$
$$=\int_0^\infty\cdots\int_0^\infty\left[\sum_{i=1}^n\left(\frac{1}{\lambda}-X_i\right)\right]^2\cdot\lambda^n\mathrm{e}^{-\lambda(X_1+\cdots+X_n)}\mathrm{d}X_1\cdots\mathrm{d}X_n=\frac{n}{\lambda^2}.$$

故根据式 (13.2.33) 或式 (13.2.22), 均有

$$\pi(\lambda)\propto 1/\lambda.$$

显然, 这是一个广义先验分布. 当未知参数选为 $\mu=1/\lambda$(指数分布的期望值), 经过类似运算可得

$$\pi(\mu)\propto 1/\mu.$$

这与指数总体期望值 μ 的不变先验分布相同. 在此例中我们可以看到, 用式 (13.2.31) 和式 (13.2.33) 计算先验分布运算最为简便.

例 13.5　二项分布参数的 Jeffreys 先验分布

定义

$$l\equiv\ln b(r;n,p)=\ln\begin{pmatrix}n\\r\end{pmatrix}+r\ln p+(n-r)\ln(1-p),$$

因为

$$\frac{\partial^2 l}{\partial p^2}=\frac{2rp-r-np^2}{p^2(1-p)^2},$$

由式 (13.2.31) 得知

$$I_1(p)=-E\left\{\frac{\partial^2 l}{\partial p^2}\right\}=\frac{1}{p^2(1-p)^2}\sum_r(np^2+r-2rp)b(r;n,p),$$

注意到

$$\sum_r b(r;n,p)=1,\quad \sum_r rb(r;n,p)=E\left(r\right)=np,$$

故有

$$I_1(p)=\frac{1}{p^2(1-p)^2}\left[np^2+(1-2p)\sum_r b(r;n,p)\right]$$
$$=\frac{1}{p^2(1-p)^2}\left[np^2+np\,(1-2p)\right]=\frac{n}{p\,(1-p)}.$$

据式 (13.2.33) 即得先验分布为

$$\pi(p)\propto|I_1(p)|^{1/2}=[p(1-p)]^{-1/2},\quad 0<p<1.$$

例 13.6　泊松分布参数的 Jeffreys 先验分布

定义 $l\equiv\ln P(X|\mu)=X\ln\mu-\ln X!-\mu$,

注意到

$$\sum_{x=0}^{\infty}xP\left(X|\mu\right)=E\left[P\left(X|\mu\right)\right]=\mu.$$

由式 (13.2.31) 得

$$I_1(\mu)=-E\left[\frac{\partial^2 l}{\partial\mu^2}\right]=\sum_{X=0}^{\infty}\frac{X}{\mu^2}P(X|\mu)=\frac{1}{\mu}.$$

利用式 (13.2.33) 即得先验分布

$$\pi(\mu)\propto|I_1(\mu)|^{1/2}=\mu^{-1/2}.$$

例 13.7　正态分布参数的 Jeffreys 先验分布

(a) $N(\mu,\sigma)$ 中参数 μ 的先验分布 $\pi(\mu)$(σ 为已知常数).

定义

$$l\equiv\ln N(\mu,\sigma)=-\ln(\sqrt{2\pi}\sigma)-\frac{1}{2\sigma^2}(X-\mu)^2,$$

因为

$$I_1(\mu)=-E\left\{\frac{\partial^2 l}{\partial\mu^2}\right\}=\frac{1}{\sigma^2},$$

据式 (13.2.33) 得

$$\pi(\mu)\propto 1.$$

这与正态总体不变先验分布中的位置参数先验分布相同.

(b) 正态总体 $N(\mu,\sigma^2)$ 中参数 σ 和 $\delta=\sigma^2$ 的先验分布 $\pi(\sigma)$ 和 $\pi(\delta)$ (μ 为已知常数).

令

$$l\equiv\ln N(\mu,\sigma^2)=-\ln(\sqrt{2\pi}\sigma)-\frac{(X-\mu)^2}{2\sigma^2}=-\frac{1}{2}\ln(2\pi\delta)-\frac{(X-\mu)^2}{2\delta},$$

由式 (13.2.31) 知

$$I_1(\sigma) = -E\left\{\frac{\partial^2 l}{\partial \sigma^2}\right\} = -E\left\{\frac{1}{\sigma^2}\left[1 - \frac{3(X-\mu)^2}{\sigma^2}\right]\right\} = \frac{2}{\sigma^2},$$

$$I_1(\delta) = -E\left\{\frac{\partial^2 l}{\partial \delta^2}\right\} = -E\left\{\frac{1}{\delta^2}\left[\frac{1}{2} - \frac{(X-\mu)^2}{\delta}\right]\right\} = \frac{1}{2\delta^2},$$

据式 (13.2.33) 得

$$\pi(\sigma) \propto \frac{1}{\sigma}, \quad \pi(\delta) \propto \frac{1}{\delta}.$$

其中 $\pi(\sigma)$ 与正态总体不变先验分布中的尺度参数先验分布相同.

(c) 正态总体 $N(\mu, \sigma^2)$ 中参数 $\boldsymbol{\vartheta} = (\mu, \sigma)^T$ 和 $\boldsymbol{\vartheta}' = (\mu, \delta)^T (\delta = \sigma^2)$ 的先验分布 $\pi(\boldsymbol{\vartheta})$ 和 $\pi(\boldsymbol{\vartheta}')$.

令

$$l \equiv \ln N(\mu, \sigma^2) = -\ln(\sqrt{2\pi}\sigma) - \frac{(X-\mu)^2}{2\sigma^2} = -\frac{1}{2}\ln(2\pi\delta) - \frac{(X-\mu)^2}{2\delta},$$

由式 (13.2.32) 得

$$I_1(\boldsymbol{\vartheta}) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix}, \quad I_1(\boldsymbol{\vartheta}') = \begin{pmatrix} \frac{1}{\delta} & 0 \\ 0 & \frac{1}{2\delta^2} \end{pmatrix},$$

故有

$$\pi(\boldsymbol{\vartheta}) \propto \frac{1}{\sigma^2}, \quad \pi(\boldsymbol{\vartheta}') \propto \frac{1}{\delta^{3/2}}.$$

其中 $\pi(\boldsymbol{\vartheta})$ 与正态总体不变先验分布中的位置－尺度参数先验分布不相同.

Jeffreys 原则的 $\pi(\boldsymbol{\vartheta}) \propto \dfrac{1}{\sigma^2}$ 表明 $\vartheta_1 = \mu$ 与 $\vartheta_2 = \sigma$ 是不独立的.

(4) **共轭分布方法**

H. Raiffa 和 R. Schlaifer[87] 提出先验分布应取共轭分布才合适.

设总体 X 的概率密度为 $f(X|\vartheta)$, 其容量 n 的样本似然函数为 $L(\boldsymbol{X}|\vartheta)$, 如果选择先验分布 $\pi(\vartheta)$ 与式 (13.2.5) 定义的后验分布 $h(\boldsymbol{X}|\vartheta)$ 属于同一分布类型 (即同一分布族), 则 $\pi(\vartheta)$ 称为参数 ϑ 的**共轭先验分布**.

由此定义可见, 共轭先验分布是对某一分布中的参数而言的, 如总体分布的均值、方差等等. 脱离指定参数及其分布谈论共轭先验分布是没有意义的.

文献 [35] 表明, 若先验分布为离散分布, 则它一定是共轭分布；若先验分布为连续分布, 则可以证明仅当总体密度函数 $f(X|\vartheta)$ 具有指数形式才存在共轭分布, 这时 $\pi(\vartheta)$ 具有如下形式

$$\pi(\vartheta) \propto G(\vartheta)^\alpha \exp\{\beta A(\vartheta)\}, \tag{13.2.35}$$

其中 α, β 与 ϑ 无关, A, G 为 ϑ 的任意函数.

共轭分布要求先验分布 $\pi(\vartheta)$ 与后验分布 $h(\boldsymbol{X}|\vartheta)$ 属于同一分布族, 注意到

$$h(\vartheta|\boldsymbol{X}) \propto \pi(\vartheta)L(\vartheta|\boldsymbol{X}),$$

可见共轭分布要求先验分布 $\pi(\vartheta)$ 提供的信息与样本分布 $L(\boldsymbol{X}|\vartheta)$ 提供的信息综合以后, 不改变 ϑ 的总的分布规律.

共轭分布要求先验分布与后验分布属于同一分布族, 这实质是要求先前的经验知识 (先验分布) 和样本提供的信息具有某种同一性, 要求过去的经验知识通过综合样本信息后能转化为同一类型的经验知识. 这一后验分布可以作为进一步的试验或观察的先验分布. 这样人们对 ϑ 的认识就能不断深化.

共轭分布要求先验分布与后验分布属于同一分布族还意味着, 先验分布 $\pi(\vartheta)$ 与后验分布 $h(\boldsymbol{X}|\vartheta)$ 的差别仅仅在于先验分布中的参数需要更新, 即先验分布写成 $\pi(\vartheta;\alpha)$, 则后验分布可表示为 $\pi(\vartheta;\beta)$, 其中 β 为 α 和样本观测值 $\boldsymbol{X}$ 的函数 $\beta \equiv \beta(\alpha, \boldsymbol{X})$:

$$\pi(\vartheta;\beta) = \frac{L(\boldsymbol{X}|\vartheta)\pi(\vartheta;\alpha)}{\displaystyle\int L(\boldsymbol{X}|\vartheta)\pi(\vartheta;\alpha)\mathrm{d}\vartheta}. \tag{13.2.36}$$

根据共轭分布方法中先验分布和样本信息具有同一性这一性质, 在许多情形下, 共轭先验分布可以用下述方法获得: 首先求出似然函数 $L(\boldsymbol{X}|\vartheta)$, 按照式 (13.2.9), 根据 $L(\boldsymbol{X}|\vartheta)$ 中所含 ϑ 的因式情况, 选取与似然函数 (作为随机变量 ϑ 的函数) 具有相同核的分布作为先验分布, 这个分布往往就是共轭先验分布.

下面我们用几种常见的总体分布作为例子来说明这一方法.

例 13.8 二项分布参数的共轭先验分布

设总体 X 服从二项分布 $b(n,p)$, 容量 n 的样本的似然函数为

$$L(X|p) = \begin{pmatrix} n \\ X \end{pmatrix} p^X(1-p)^{n-X} \propto p^X(1-p)^{n-X}.$$

$L(X|p)$ 中所含 p 的因式为贝塔 (β) 分布的核, 因贝塔分布的概率密度为

$$\beta(p;a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{a-1}(1-p)^{b-1}, \quad 0 \leqslant p \leqslant 1,$$

a, b 是贝塔分布的参数 (正整数). 从而可设 p 的先验分布为贝塔分布 $\beta(p;a,b)$, 其核为 $p^{a-1}(1-p)^{b-1}$, 代入式 (13.2.9) 立即得到后验分布

$$h(p;a,b|X) \propto p^{X+a-1}(1-p)^{n-X+b-1} \sim \beta(X+a, n-X+b).$$

因此, 先验分布与后验分布同为贝塔分布.

例 13.9　几何分布参数的共轭先验分布

设总体 X 服从几何分布, 概率密度为 $f(X|p) = p(1-p)^{X-1}, p \in [0,1]$, $X = 1, \cdots$, 总体容量 n 的样本的似然函数为

$$L(\boldsymbol{X}|p) = \prod_{i=1}^{n} p(1-p)^{X_i-1} = p^n(1-p)^{n\left(\bar{X}-1\right)},$$

式中 $\bar{X}$ 是样本平均. $L(\boldsymbol{X}|p)$ 中所含 p 的因式为贝塔分布的核, 从而可设 p 的先验分布为贝塔分布 $\beta(p;a,b)$. 利用式 (13.2.9) 求得后验分布为

$$h(p;a,b|\boldsymbol{X}) \propto p^{n+a-1}(1-p)^{b+n\left(\bar{X}-1\right)-1},$$

忽略常数乘因子, $h(p;a,b|\boldsymbol{X})$ 正是参数为 $n+a$ 和 $n\left(\bar{X}-1\right)+b$ 的贝塔分布：

$$h(p;a,b|\boldsymbol{X}) \sim \beta\left(n+a, n\left(\bar{X}-1\right)+b\right).$$

因此先验分布和后验分布同属一类分布, 贝塔分布是几何分布总体的共轭先验分布族. 先验分布的参数 a,b 在后验分布中更新为 $n+a$ 和 $n\bar{X}+b$, 它们是先验分布的参数 a,b 和样本观测值 $n,\bar{X}$ 的函数.

例 13.10　泊松分布参数的共轭先验分布

设总体 X 服从泊松分布 $P(X|\mu) = \dfrac{1}{X!}\mu^X \mathrm{e}^{-\mu}, X = 0,1,\cdots$. 总体容量 n 的样本的似然函数为

$$L(\boldsymbol{X}|\mu) = \prod_{i=1}^{n} \frac{1}{X_i!}\mu^{X_i}\mathrm{e}^{-\mu} = \frac{\mathrm{e}^{-n\mu}\mu^{n\bar{X}}}{X_1!\cdots X_n!} \propto \mathrm{e}^{-n\mu}\mu^{n\bar{X}},$$

式中 $\bar{X}$ 是样本平均. 由于伽马 $(\varGamma)$ 分布为

$$\varGamma(\mu;a,b) = \frac{b^a}{\varGamma(a)}\mu^{a-1}\mathrm{e}^{-b\mu}, \quad a,b>0, \quad 0 \leqslant \mu < \infty.$$

$L(\boldsymbol{X}|\mu)$ 中所含 μ 的因式为伽马 $(\varGamma)$ 分布关于参数 μ 的核, 因而可选择伽马分布 $\varGamma(\mu;a,b)$ 为参数 μ 的先验分布, a,b 是伽马分布的参数, 忽略常数乘因子有：

$$\pi(\mu;a,b) = \varGamma(\mu;a,b) \propto \mu^{a-1}\mathrm{e}^{-b\mu}, \quad a,b>0, \quad 0 \leqslant \mu < \infty.$$

利用式 (13.2.9) 求得后验分布为

$$h(\mu|\boldsymbol{X}) \propto \mu^{n\bar{X}+a-1}\mathrm{e}^{-\mu(n+b)},$$

即有 $h(\mu|\mathrm{X}) \sim \varGamma(\mu; a+n\bar{X}, b+n)$, 因此先验分布和后验分布同属伽马分布, 伽马分布是泊松总体的共轭先验分布族.

例 13.6 中用 Jeffreys 原则算出的泊松总体先验分布 $\pi(\mu) \propto 1/\mu^{1/2}$ 可以视为 $a=1/2, b=0$ 的伽马分布, 其后验分布由上式立即得到:

$$h(\mu; a, b|\boldsymbol{X}) \sim \Gamma\left(\mu; n\bar{X}+\frac{1}{2}, n\right) \propto \mu^{n\bar{X}-\frac{1}{2}}\mathrm{e}^{-\mu n}.$$

例 13.11　指数分布参数的共轭先验分布

设总体 X 的概率密度为 $f(X|\lambda)=\lambda\exp(-\lambda X), X>0$. 总体容量 n 的样本的似然函数为

$$L(\boldsymbol{X}|\lambda)=\prod_{i=1}^{n}\lambda\mathrm{e}^{-\lambda X_i}=\lambda^n\mathrm{e}^{-\lambda n\bar{X}},$$

式中 $\bar{X}$ 是样本平均. $L(\boldsymbol{X}|\lambda)$ 中所含 λ 的因式为伽马分布的核, 因而可选择伽马分布 $\Gamma(\lambda; a, b)$ 为参数 λ 的先验分布, 利用式 (13.2.9) 求得后验分布为

$$h(\lambda; a, b|\boldsymbol{X}) \propto \lambda^{n+a-1}\mathrm{e}^{-\lambda(n\bar{X}+b)},$$

忽略常数乘因子, $h(\lambda; a, b|\boldsymbol{X})$ 正是参数为 $n+a$ 和 $n\bar{X}+b$ 的伽马分布:

$$h(\lambda; a, b|\boldsymbol{X}) \sim \Gamma(\lambda; n+a, n\bar{X}+b).$$

因此先验分布和后验分布同属一类分布, 伽马分布是指数总体的共轭先验分布族. 先验分布的参数 a, b 在后验分布中更新为 $n+a$ 和 $n\bar{X}+b$, 它们是先验分布的参数 a, b 和样本观测值 $n, \bar{X}$ 的函数.

例 13.4 中用 Jeffreys 原则算出的指数总体先验分布 $\pi(\lambda) \propto \lambda^{-1}$ 可以视为 $a=0, b=0$ 的伽马分布, 其后验分布由上式立即得到:

$$h(\lambda; 0, 0|\boldsymbol{X}) \sim \Gamma(\lambda; n, n\bar{X}) \propto \lambda^{n-1}\mathrm{e}^{-\lambda n\bar{X}}.$$

如设总体 X 的概率密度为 $f(X|\vartheta)=\vartheta^{-1}\exp(-X/\vartheta), X>0$. 总体容量 n 的样本的似然函数为

$$L(\boldsymbol{X}|\vartheta)=\prod_{i=1}^{n}\vartheta^{-1}\mathrm{e}^{-X_i/\vartheta}=\vartheta^{-n}\mathrm{e}^{-n\bar{X}/\vartheta}.$$

$L(\boldsymbol{X}|\vartheta)$ 中所含 ϑ 的因式为倒伽马分布的核 (见示例 13.12 式 (13.2.37) 所示的倒伽马分布密度函数), 因而可选择倒伽马分布 $I\Gamma(\vartheta; a, b)$ 为参数 ϑ 的先验分布, 利用式 (13.2.9) 求得后验分布为

$$h(\vartheta; a, b|\boldsymbol{X}) \propto \vartheta^{-(a+1+n)}\mathrm{e}^{-(n\bar{X}+b)/\vartheta},$$

忽略常数乘因子, $h(\vartheta;a,b|\boldsymbol{X})$ 正是参数为 $n+a$ 和 $n\bar{X}+b$ 的倒伽马分布：

$$h(\vartheta;a,b|\boldsymbol{X}) \sim I\Gamma(\vartheta;n+a,n\bar{X}+b).$$

因此先验分布和后验分布同属一类分布.

例 13.4 中用 Jeffreys 原则算出的指数总体先验分布 $\pi(\vartheta) \propto \vartheta^{-1}$ 可以视为 $a=0,b=0$ 的倒伽马分布, 其后验分布由上式立即得到：

$$h(\vartheta;0,0|\boldsymbol{X}) \sim I\Gamma(\vartheta;n,n\bar{X}) \propto \vartheta^{-(n+1)}\mathrm{e}^{-n\bar{X}/\vartheta}.$$

例 13.12　正态分布参数的共轭先验分布

(a) 求正态总体 $N(\mu,\sigma^2)$ 参数 μ 的共轭先验分布. (σ^2 已知)

总体容量 n 的样本的似然函数为

$$L(\boldsymbol{X}|\mu)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(X_i-\mu)^2}{2\sigma^2}} \propto \mathrm{e}^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu)^2},$$

当仅把 μ 视为变量, 上式右边求和号内的项 $\sum\limits_{i=1}^{n}(X_i-\mu)^2$ 总可以写成 $n(\mu-a)^2+b$ 的形式, 其中 a,b 与 μ 无关, 略去与 μ 无关的常数项, 有

$$L(\boldsymbol{X}|\mu) \propto \mathrm{e}^{-c(\mu-a)^2},$$

其中 c 与 μ 无关, 即似然函数具有变量 μ 的正态分布的核. 因此可选择正态分布为参数 μ 的先验分布 $N(\mu_0,\tau^2)$, 即

$$\pi(\mu) \propto \exp\left[-\frac{(\mu-\mu_0)^2}{2\tau^2}\right].$$

利用式 (13.2.9) 求得后验分布为

$$h(\mu|\boldsymbol{X}) \propto \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2-\frac{(\mu-\mu_0)^2}{2\tau^2}\right] \propto \exp\left[-\frac{(\mu-\mu')^2}{2\gamma^2}\right],$$

其中,

$$\mu'=(\bar{X}n/\sigma^2+\mu_0/\tau^2)/(n/\sigma^2+1/\tau^2),$$

$$\gamma^2=1/(n/\sigma^2+1/\tau^2)=\frac{\sigma^2\tau^2}{n\tau^2+\sigma^2}.$$

即有 $h(\mu|\boldsymbol{X}) \sim N(\mu',\gamma^2)$, 因此先验分布和后验分布同属一类分布, 正态分布是正态分布总体 $N(\mu,\sigma^2)$ 的共轭先验分布族.

当 $\tau\to\infty,\mu'\to\bar{X},\gamma^2\to\sigma^2/n$, $h(\mu|\boldsymbol{X})\sim N(\bar{X},\sigma^2/n)$, 这时后验概率完全由样本的性质决定, 这相当于对先验信息一无所知的情形.

(b) 求正态总体 $N(\mu,\sigma^2)$ 参数 $\delta=\sigma^2$ 的共轭先验分布. (μ 已知)

总体容量 n 的样本的似然函数为

$$L(\boldsymbol{X}|\mu)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(X_i-\mu)^2}{2\sigma^2}}\propto\left(\frac{1}{\delta}\right)^{n/2}\mathrm{e}^{-\frac{1}{2\delta}\sum\limits_{i=1}^{n}(X_i-\mu)^2},$$

该似然函数中所含 δ 的因式将决定 $\delta=\sigma^2$ 的共轭先验分布的形式. 什么分布具有上述函数的核呢?

设变量 Y 服从伽马分布 $\Gamma(a,b)$, 通过概率运算可以求得变量 $X=Y^{-1}$ 服从**倒伽马分布**$I\Gamma(X;a,b)$, 其概率密度为

$$I\Gamma(X;a,b)=\frac{b^a}{\Gamma(a)}\left(\frac{1}{X}\right)^{a+1}\mathrm{e}^{-\frac{b}{X}},\quad a,b>0,\quad X>0. \tag{13.2.37}$$

$I\Gamma(X;a,b)$ 分布期望值和方差为:

$$E\left(X\right)=\frac{b}{a-1},\quad V\left(X\right)=\frac{b^2}{\left(a-1\right)^2\left(a-2\right)} \tag{13.2.38}$$

与似然函数对比可知似然函数具有倒伽马分布的核, 因此 $\delta=\sigma^2$ 的先验分布可取为倒伽马分布 $I\Gamma(a,b)$. 利用式 (13.2.9) 求得后验分布为

$$h(\delta;a,b|\boldsymbol{X})\propto\left(\frac{1}{\delta}\right)^{\frac{n}{2}+a+1}\exp\left[-\frac{1}{\delta}\left(b+\frac{1}{2}\sum_{i=1}^{n}(X_i-\mu)^2\right)\right],$$

忽略常数乘因子, $h(\delta;a,b|\boldsymbol{X})$ 正是参数为 $n/2+a$ 和 $b+\dfrac{1}{2}\sum\limits_{i=1}^{n}(X_i-\mu)^2$ 的倒伽马分布:

$$h(\delta;a,b|\boldsymbol{X})\sim I\Gamma\left\{\delta;\frac{n}{2}+a,b+\frac{1}{2}\sum_{i=1}^{n}(X_i-\mu)^2\right\}.$$

因此先验分布和后验分布同属一类分布, 伽马分布是指数总体的共轭先验分布族. 先验分布的参数 a,b 在后验分布中更新为 $n/2+a$ 和 $b+\sum\limits_{i=1}^{n}(X_i-\mu)^2/2$, 它们是先验分布的参数 a,b 和样本观测值的函数.

(c) 求正态总体 $N(\mu,\sigma^2)$ 参数 $\vartheta=(\vartheta_1,\vartheta_2)^T=\left(\mu,\delta=\sigma^2\right)^T$ 的共轭先验分布. (μ 和 δ 均未知).

总体容量 n 的样本的似然函数为

$$L(\boldsymbol{X}|\vartheta)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\delta}}\mathrm{e}^{-\frac{(X_i-\mu)^2}{2\delta}}\propto\left(\frac{1}{\delta}\right)^{n/2}\mathrm{e}^{-\frac{z}{2\delta}-\frac{n\left(\bar{X}-\mu\right)^2}{2\delta}},$$

其中,

$$z\equiv\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$$

该似然函数中所含 μ 和 δ 的因式将决定 $\delta=\sigma^2$ 的共轭先验分布的形式. 什么分布具有上述函数的核呢?

设 δ 的先验分布为 $I\Gamma(a,\lambda)$

$$\pi\left(\delta\right)\propto\delta^{-(a+1)}\mathrm{e}^{-\frac{\lambda}{\delta}},$$

又设在给定 δ 的条件下, μ 的条件密度为

$$N\left(\mu_0,\frac{\delta}{\tau}\right)=\frac{\tau}{\sqrt{2\pi\delta}}\exp\left[-\frac{\tau}{2\delta}\left(\mu-\mu_0\right)^2\right],$$

于是 $\vartheta=\left(\mu,\delta\right)^T$ 的密度为

$$\pi\left(\vartheta\right)\propto\delta^{-\left(a+\frac{3}{2}\right)}\mathrm{e}^{-\frac{\lambda}{\delta}}\mathrm{e}^{-\frac{\tau(\mu-\mu_0)^2}{2\delta}},$$

这一分布记为

$$G\left(a,\lambda;\mu_0,\tau\right),\quad a,\lambda,\tau>0,\quad \mu_0\in(-\infty,\infty).$$

故有

$$\pi\left(\vartheta\right)\sim G\left(a,\lambda;\mu_0,\tau\right)\propto\delta^{-\left(a+\frac{3}{2}\right)}\mathrm{e}^{-\frac{\lambda}{\delta}}\mathrm{e}^{-\frac{\tau(\mu-\mu_0)^2}{2\delta}},$$

利用式 (13.2.9) 求得后验分布为

$$\begin{aligned}h(\vartheta|\boldsymbol{X})=h(\mu,\delta|\boldsymbol{X})&\propto\delta^{-\left(\frac{n+3}{2}+a\right)}\mathrm{e}^{-\frac{\lambda}{\delta}}\mathrm{e}^{-\frac{\tau(\mu-\mu_0)^2+n\left(\bar{X}-\mu\right)^2+z}{2\delta}}\\&\propto\delta^{-\left(\frac{n+3}{2}+a\right)}\exp\left\{-\left[\lambda+\frac{z}{2}+\frac{n\tau}{2\left(n+\tau\right)}\left(\bar{X}-\mu_0\right)^2\right]\Big/\delta\right\}\\&\quad\cdot\exp\left[-\frac{\left(n+\tau\right)\left(\mu-\hat{\mu}\right)^2}{2\delta}\right]\\&\sim G\left(a+\frac{n}{2},\lambda+\frac{z}{2}+\frac{n\tau}{2\left(n+\tau\right)}\left(\bar{X}-\mu_0\right)^2;\hat{\mu},n+\tau\right),\end{aligned}$$

其中,

$$\hat{\mu}=\frac{n\bar{X}+\tau\mu_0}{n+\tau},$$

因而

$$\tau\left(\mu-\mu_0\right)^2+n\left(\bar{X}-\mu\right)^2=(n+\tau)\left(\mu-\hat{\mu}\right)^2+\frac{n\tau}{n+\tau}\left(\bar{X}-\mu_0\right)^2.$$

因此先验分布和后验分布同属一类分布.

由以上例子可以看到, 由于共轭先验分布与样本的似然函数 $L(\boldsymbol{X}|\vartheta)$ 属于同一分布族, 故当先验分布选取共轭先验分布的情形下后验分布的确定特别简单. 先验分布中所含的参数称为**超参数**或**冗余参数**(nuisance parameter), 共轭先验分布常含有超参数, 恰当地确定超参数是选用共轭先验分布需解决的问题. 这一问题的解决没有通用的方案, 取决于已有的先验信息的种类和多少. 下面就选用贝塔分布 $\beta(a,b)$ 作为共轭先验分布的情形讨论其超参数 a 和 b 的确定.

(a) 利用先验矩

假定根据先验信息已获得参数 ϑ 的若干估计值 $\vartheta_1,\cdots,\vartheta_k$, 据此可得先验均值 $\bar{\vartheta}$ 和先验方差 S_ϑ^2：

$$\bar{\vartheta}=\frac{1}{k}\sum_{i=1}^{k}\vartheta_i,\quad S_\vartheta^2=\frac{1}{k-1}\sum_{i=1}^{k}\left(\vartheta_i-\bar{\vartheta}\right)^2.$$

将它们作为贝塔分布 $\beta(a,b)$ 的期望值和方差的估计即得

$$\begin{cases}\dfrac{a}{a+b}=\bar{\vartheta},\\[2mm] \dfrac{ab}{(a+b)^2(a+b+1)}=S_\vartheta^2.\end{cases}$$

解此方程组, 即得超参数 a 和 b 的矩法估计：

$$\begin{cases}\hat{a}=\bar{\vartheta}\left[\dfrac{\left(1-\bar{\vartheta}\right)\bar{\vartheta}}{S_\vartheta^2}-1\right],\\[2mm] \hat{b}=\left(1-\bar{\vartheta}\right)\left[\dfrac{\left(1-\bar{\vartheta}\right)\bar{\vartheta}}{S_\vartheta^2}-1\right].\end{cases}$$

(b) 利用先验分位数

假定根据先验信息已知获得参数 ϑ 的先验贝塔分布 $\beta(a,b)$ 的上下四分位数 ϑ_U 和 ϑ_L, 根据贝塔分布性质, 即可列出方程组

$$\begin{cases}\displaystyle\int_0^{\vartheta_L}\frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}\vartheta^{a-1}(1-\vartheta)^{b-1}\,\mathrm{d}\vartheta=0.25,\\[2mm] \displaystyle\int_{\vartheta_U}^{1}\frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}\vartheta^{a-1}(1-\vartheta)^{b-1}\,\mathrm{d}\vartheta=0.25.\end{cases}$$

利用数值积分方法求解该方程组可得超参数 a 和 b 的估计值.

13.3.2 节讨论的经验贝叶斯方法中, 阐述了利用历史样本值估计参数 ϑ 先验分布 $\pi(\vartheta)$ 中的超参数的方法.

(5) **选择先验分布方法的小结和讨论**

前面讨论了选择待定参数ϑ 的先验分布的几种方法, 即贝叶斯假设、不变先验分布、Jeffreys 原则和共轭先验分布. 在实际问题中究竟选择什么样的先验分布是恰当的, 必须视具体情况而定.

表 13.2 列出了几种经常遇到的总体的待定参数ϑ 用以上四种方法确定的先验分布 $\pi(\vartheta)$ 的核. 需要指出的是共轭先验分布只是提供了参数 ϑ 的先验分布的函数类型, 其中的超参数 (a,b,μ,τ,λ) 需要适当地加以确定.

有了先验分布, 即可根据式 (13.2.8) 或式 (13.2.9) 计算后验分布. 在许多情形下, 由贝叶斯假设、不变性原理和 Jeffreys 原则确定的先验分布可视为共轭先验分布的特例, 例如贝叶斯假设确定的先验分布可表示为 (或形式地表示为)$\pi(\vartheta) = 1 = \beta(1,1) = \Gamma(1,0) = I\Gamma(-1,0)$, 先验分布 $1/\vartheta$ 可表示为 (或形式地表示为) $\pi(\vartheta) = 1/\vartheta = \beta(0,1) = \Gamma(0,0) = I\Gamma(0,0)$. 这样, 根据共轭分布法, 不需要进行繁杂的计算即可求得相应的后验分布. 不同的先验分布对应的后验分布亦列于表 13.2.

如果在随机试验之前对于参数 ϑ 基本无所了解, 无历史资料可查, 无先前的经验可借鉴, 显然, 无信息先验分布是比较合理的选择. 贝叶斯假设、不变先验分布、Jeffreys 原则都属于无信息先验分布, 但是对同样的总体会给出不同的先验分布. 贝叶斯假设最为简单, 但在许多情况下不够合理. 对于总体分布参数属于位置–尺度族的情形, 不变先验分布是比较适当的选择. Jeffreys 原则是一种更广泛意义上的不变原则, 而且利用费歇尔信息量, 容易从总体的概率密度求得参数 ϑ 的先验分布, 因此, 它成为最广泛使用的无信息先验分布.

如果在随机试验之前对于参数 ϑ 的分布已有一定了解, 有过去的试验结果可查, 有先前的经验可借鉴, 也就是说, 对参数 ϑ 的先验分布的类型、特征以及它的某些数字特征如期望值、方差、分位数等有一定的知识, 即使不是那么确定, 这些以往的知识对于确定参数 ϑ 的先验分布也是有帮助的, 应当加以利用. 在这种情况下, 共轭先验分布是一种比较合理的选择. 需要注意的是共轭先验分布只是提供了参数 ϑ 的先验分布的函数类型, 其中的超参数 (例如表 13.2 中的 a,b,μ,τ,λ) 需要利用以往的知识适当地加以确定.

先验分布的合理性和正确性问题是贝叶斯理论中的一大难点, 也是受到其他学派批评最多的主要问题. 本节中对于先验分布选择方法的介绍和讨论只是帮助读者对这一问题的现状有所了解. 先验分布的合理性和正确性问题还需要进一步探讨和研究.

表 13.2 不同总体的先验分布 $\pi(\vartheta)$ 和后验分布 $h(\vartheta|X)$ 的核

总体分布 \ 方法		贝叶斯假设	不变性原则	Jeffreys 原则	共轭分布法
二项分布	$\binom{n}{X}\vartheta^X(1-\vartheta)^{n-X}$	1		$[\vartheta(1-\vartheta)]^{-1/2}$	$\beta(a,b)\propto\vartheta^{a-1}(1-\vartheta)^{b-1}$
		$\beta(1+X,1+n-X)$		$\beta\left(\frac{1}{2}+X,\frac{1}{2}+n-X\right)$	$\beta(a+X,b+n-X)$
几何分布	$\vartheta(1-\vartheta)^{X-1}$	1		$\vartheta^{-1}(1-\vartheta)^{-1/2}$	$\beta(a,b)\propto\vartheta^{a-1}(1-\vartheta)^{b-1}$
		$\beta\left(1+n,1+n\bar{X}-n\right)$		$\beta\left(n,\frac{1}{2}+n\bar{X}-n\right)$	$\beta\left(a+n,b+n\bar{X}-n\right)$
泊松分布	$\frac{1}{X!}\vartheta^X\mathrm{e}^{-\vartheta}$	1		$1/\vartheta^{1/2}$	$\Gamma(a,b)\propto\vartheta^{a-1}\mathrm{e}^{-b\vartheta}$
		$\Gamma(1+n\bar{X},n)$		$\Gamma\left(\frac{1}{2}+n\bar{X},n\right)$	$\Gamma(a+n\bar{X},b+n)$
指数分布	$\vartheta\mathrm{e}^{-\vartheta X}$	1		$1/\vartheta$	$\Gamma(a,b)\propto\vartheta^{a-1}\mathrm{e}^{-b\vartheta}$
		$\Gamma(1+n,n\bar{X})$		$\Gamma(n,n\bar{X})$	$\Gamma(a+n,b+n\bar{X})$
	$\frac{1}{\vartheta}\mathrm{e}^{-X/\vartheta}$	1	$1/\vartheta$	$1/\vartheta$	$I\Gamma(a,b)\propto\left(\frac{1}{\vartheta}\right)^{a+1}\mathrm{e}^{-\frac{b}{\vartheta}}$
		$I\Gamma(n-1,n\bar{X})$	同右	$I\Gamma(n,n\bar{X})$	$I\Gamma(a+n,b+n\bar{X})$

续表

总体分布 \ 方法		贝叶斯假设	不变性原则	Jeffreys 原则	共轭分布法
正态分布	$N(\vartheta,\sigma^2)$	1	1	1	$N(\mu,\tau^2)$
		$N\left(\bar{X},\dfrac{\sigma^2}{n}\right)$	同左	同左	$N(\hat{\mu},\gamma^2),\left(\hat{\mu}=\dfrac{\dfrac{n\bar{X}}{\sigma^2}+\dfrac{\mu}{\tau^2}}{\dfrac{n}{\sigma^2}+\dfrac{1}{\tau^2}},\quad \gamma^2=\dfrac{\sigma^2\tau^2}{n\tau^2+\sigma^2}\right)$
	$N(\mu_0,\vartheta^2)$	1	$1/\vartheta$	$1/\vartheta$	
	$N(\mu_0,\vartheta)$, $\vartheta=\sigma^2$	1		$1/\vartheta$	$I\Gamma(a,b)\propto\left(\frac{1}{\vartheta}\right)^{a+1}\mathrm{e}^{-\frac{b}{\vartheta}}$
				$I\Gamma\left(\dfrac{n}{2},\dfrac{1}{2}\sum\limits_{i=1}^{n}(X_i-\mu_0)^2\right)$	$I\Gamma\left(a+\dfrac{n}{2},b+\dfrac{1}{2}\sum\limits_{i=1}^{n}(X_i-\mu_0)^2\right)$
	$N(\vartheta_1,\vartheta_2^2)$	1	$1/\vartheta_2$	$1/\vartheta_2^2$	
	$N(\vartheta_1,\vartheta_2)$, $\vartheta_2=\sigma^2$	1		$1\big/\vartheta_2^{3/2}$	$G(a,\lambda;\mu,\tau)\propto$ $\vartheta_2^{-\left(a+\frac{3}{2}\right)}\mathrm{e}^{-\frac{\lambda}{\vartheta_2}}\mathrm{e}^{-\frac{\tau(\vartheta_1-\mu)^2}{2\vartheta_2}}$
		$G\left(\dfrac{n-3}{2},\dfrac{z}{2};\bar{X},n\right)$ 其中 $z=\sum\limits_{i=1}^{n}\left(X_i-\bar{X}\right)^2$		$G\left(\dfrac{n}{2},\dfrac{z}{2};\bar{X},n\right)$ 其中 $z=\sum\limits_{i=1}^{n}\left(X_i-\bar{X}\right)^2$	$G\left(a+\dfrac{n}{2},\lambda+\dfrac{z}{2}+\dfrac{n\tau\left(\bar{X}-\mu\right)^2}{2\left(n+\tau\right)};\hat{\mu},\tau+n\right)$ 其中 $z=\sum\limits_{i=1}^{n}\left(X_i-\bar{X}\right)^2,\quad \hat{\mu}=\dfrac{n\bar{X}+\mu\tau}{n+\tau}$.

表中, 对每种总体分布, $\pi(\vartheta)$ 写在上部格子中, $h(\vartheta|\boldsymbol{X})$ 写在下部格子中. ϑ 或 ϑ_i 为待估计参数,X 为测量值, $\bar{X}$ 和 n 为样本平均和样本容量, a,b,μ,τ,λ 为超参数

13.3 贝叶斯统计推断

假定我们有概率密度为 $f(X,\vartheta)$ 的随机变量 X 的容量 n 的样本观测值 $\boldsymbol{X}=(X_1,\cdots,X_n)^T$, 其似然函数为 $L(\boldsymbol{X}|\vartheta)$, 其未知参数 ϑ 的先验分布已选定为 $\pi(\vartheta)$, 并根据式 (13.2.5) 推得了相应的后验分布 $h(\vartheta|\boldsymbol{X})$. 我们需要解决的问题是根据这些知识对总体分布的参数 ϑ 的数值或总体概率分布的形式进行统计推断, 统计推断的基本问题是参数估计和假设检验.

在经典统计学中, 样本是随机变量, 样本观测值 $\boldsymbol{X}$ 只是其无限次可能的试验结果的一个具体实现, 参数 ϑ 则是一个未知的待定常数；而贝叶斯学派则认为样本观测值 $\boldsymbol{X}$ 只是一组确定的数值, 没有必要放在无限多可能值之一的地位上考察, 而参数 ϑ 则是随机变量, 贝叶斯统计推断的任务是基于已知的样本观测值 $\boldsymbol{X}$ 对未知的随机变量 ϑ 根据后验分布 $h(\vartheta|\boldsymbol{X})$ 做出推断.

贝叶斯统计推断原则认为, 对于参数 ϑ 所做的任何统计推断必须而且只能基于 ϑ 的后验分布.

13.3.1 统计决策的基本概念

参数估计和假设检验可以看作一种特殊的统计决策问题, 为此对统计决策理论的若干基本概念做一初步介绍.

(1) 统计决策问题三要素

为了估计一个未知参数, 需要确定一个适当的估计量, 该估计量就是该统计推断问题的解. 一般地说, 一个统计推断问题的解就是所谓的统计决策函数. 构成一个统计决策问题的三要素是：(a) 样本空间和分布族, (b) 决策空间, (c) 损失函数.

(a) 样本空间和分布族

设总体 X 的分布函数为 $F(x;\vartheta)$, ϑ 是未知参数 $\vartheta\in\Theta$, Θ 为参数空间. $(X_1,\cdots,X_n)$ 为总体 X 的样本, 则样本所有可能值的集合为样本空间, 记为 Ω_X. $(X_1,\cdots,X_n)$ 的联合分布函数为

$$F(x_1,\cdots,x_n;\vartheta)=\prod_{i=1}^{n}F(x_i;\vartheta),\quad \vartheta\in\Theta.$$

记 $F^*=\left\{\prod\limits_{i=1}^{n}F(x_i;\vartheta),\vartheta\in\Theta\right\}$, 则称 F^* 为样本 $(X_1,\cdots,X_n)$ 的分布族.

(b) 决策空间

在统计决策中对一个具体问题的回答称为一个**决策**, 如对参数 ϑ 的点估计问题, 一个具体的估计值就是一个决策.

一个统计决策问题中可能选取的全部决策的集合称为**决策空间**, 记为 $\mathcal{D}$. 一个决策空间至少应包含两个决策.

例如若正态分布 $N(\mu,\sigma^2)$的均值为待估计参数, 因为 $\mu \in \Theta = (-\infty,+\infty)$, 所以任意实数都可以作为 μ 的一个估计, 即关于 μ 的一个决策, 故决策空间为 $\mathcal{D}=(-\infty,+\infty)$.

(c) 损失函数

对于一个统计决策问题, 一般有多种决策可以采用, 决策的优劣用损失函数来衡量.

损失函数是依赖于参数值 $\vartheta\in\Theta$ 和决策 $d\in\mathcal{D}$ 的二元非负实函数 $L(\vartheta,d)\geqslant 0$, 它表示当参数真值为 ϑ 而采用决策 d 时, 由于 d 与 ϑ 的不同所造成的损失. 最小的损失值为 0. 在统计推断问题中, 总是利用有限的样本对总体进行推断, 误差不可避免, 因而总会带来损失, 这是损失函数定义为非负实函数的原因.

常见的损失函数有

$$L(\vartheta,d)=(\vartheta-d)^2 \quad (\text{平方差损失函数}), \tag{13.3.1}$$

$$L(\vartheta,d)=|\vartheta-d| \quad (\text{绝对差损失函数}), \tag{13.3.2}$$

$$L(\vartheta,d)=\begin{cases}1, & d\neq\vartheta\\ 0, & d=\vartheta\end{cases} \quad (0,1\text{损失函数}), \tag{13.3.3}$$

$$L(\vartheta,d)=\left(1-\frac{d}{\vartheta}\right)^2 \quad (\text{相对差损失函数}). \tag{13.3.4}$$

如果要对未知参数 ϑ 进行区间估计, 损失函数可取为

$$L(\vartheta,d)=(d_2-d_1), \quad \vartheta\in\Theta, d=[d_1,d_2]\in\mathcal{D}, \tag{13.3.5}$$

其中 $\mathcal{D}=\{[d_1,d_2]:-\infty<d_1<d_2<+\infty\}$. 该损失函数表示以区间长度 (d_2-d_1) 来度量决策 $d=[d_1,d_2]$ 带来的损失. 也可以取损失函数为

$$L(\vartheta,d)=\begin{cases}1, & \vartheta\notin[d_1,d_2],\\ 0, & \vartheta\in[d_1,d_2],\end{cases} \quad \vartheta\in\Theta, d=[d_1,d_2]\in\mathcal{D}, \tag{13.3.6}$$

这一损失函数表示当决策 d 正确 (即区间 $[d_1,d_2]$ 覆盖未知参数的实际值) 时, 损失为 0, 反之则损失为 1.

在实际的统计推断问题中, 选择适当的损失函数是一个关键问题. 一般来说, 选择的损失函数既要考虑适合问题的要求, 又要数学上便于处理. 其中式 (13.3.1) 的平方差损失函数是参数点估计中常用的一种损失函数.

(2) 统计决策函数及其风险函数

(a) 统计决策函数

一个统计推断问题, 现在就是在决策空间 $\mathcal{D}$ 中选取一个好的决策, 即损失较小的决策 d.

我们总是利用样本测量值 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 进行统计推断. 对应于样本空间 Ω_X 中的每一点 $\boldsymbol{X}$, 可在决策空间 $\mathcal{D}$ 中找到一点 $d(\boldsymbol{X})$ 与之对应, 这种对应关系可视为定义在样本空间 Ω_X 中而取值于决策空间 $\mathcal{D}$ 中的函数 $d(\boldsymbol{X})$, 称为**统计决策函数**. 它表示当样本值为 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 时采取决策 $d(\boldsymbol{X})=d(X_1,\cdots,X_n)$. 因此决策函数 $d(\boldsymbol{X})$ 本质上是一个统计量.

例如, 设总体 X 服从正态分布 $N(\mu,\sigma^2)$, σ^2 已知, 要求用总体 X 的样本 $(X_1,\cdots,X_n)$ 估计参数 μ(点估计). 此时可用 $d(\boldsymbol{X})=\bar{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$ 来估计参数 μ, 因此 $d(\boldsymbol{X})=\bar{X}$ 就是一个决策函数.

如果要对参数 μ 做区间估计, 可以采用决策函数

$$d(\boldsymbol{X})=\left[\bar{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\bar{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right],$$

其中 $z_{\alpha/2}$ 是正态分布的双侧 α 分位数.

(b) 风险函数

给定一个决策函数 $d(\boldsymbol{X})$ 后, 所采取的决策完全取决于样本 $\boldsymbol{X}=(X_1,\cdots,X_n)$, 因而损失必然与 $\boldsymbol{X}$ 有关, 即决策函数与损失函数 $L(\vartheta,d)$ 都是样本 $\boldsymbol{X}$ 的函数, 都是随机变量. 当样本 $\boldsymbol{X}$ 取不同的值 $\boldsymbol{x}$ 时, 对应的决策 $d(\boldsymbol{x})$ 也不相同. 因此不能用样本 $\boldsymbol{X}$ 的一次特定测量值 $\boldsymbol{x}$ 的损失 $L(\vartheta,d(\boldsymbol{x}))$ 来衡量决策的优劣, 而应该从整体上来加以评价. 为了比较决策函数的优劣, 常用的数量指标是平均损失, 即所谓的风险.

决策函数 $d(\boldsymbol{X})$ 的**风险函数**定义为

$$\begin{aligned}R(\vartheta,d)&=E_\vartheta[L(\vartheta,d(\boldsymbol{X}))]=E_\vartheta\left[L(\vartheta,d(X_1,\cdots,X_n))\right]\\&=\int_{\Omega_X}L(\vartheta,d(\boldsymbol{X}))f(\boldsymbol{X}|\vartheta)\mathrm{d}\boldsymbol{X},\end{aligned}\tag{13.3.7}$$

式中 $f(\boldsymbol{X}|\vartheta)$ 表示总体 X 的样本 $\boldsymbol{X}=(X_1,\cdots,X_n)^T$ 的联合概率密度函数, 与式 (13.2.9) 的 $L(\boldsymbol{X}|\vartheta)$ 有同样的意义, 只是为了避免与损失函数 $L(\vartheta,d)$ 混淆而采用不同的符号; E_ϑ 表示当参数为 ϑ 时, 对样本的函数 $L(\vartheta,d(\boldsymbol{X}))$ 在样本空间求数学期望. $R(\vartheta,d)$ 表示当参数真值为 ϑ 时, 采用决策 d 所蒙受的平均损失.

显然, 风险越小, 即平均损失越小, 决策函数越好. 但对于给定的决策函数 $d(\boldsymbol{X})$, 风险函数仍然是 ϑ 的函数, 所以两个决策函数风险大小的比较涉及两个函数的比较. 因此就需要有比较不同决策函数优良性的准则.

给定损失函数 $L(\vartheta, d)$, $d_1(\boldsymbol{X})$ 和 $d_2(\boldsymbol{X})$ 是统计决策问题中的两个决策函数, 若其风险函数满足不等式

$$R(\vartheta, d_1) \leqslant R(\vartheta, d_2), \quad \forall \vartheta \in \Theta, \tag{13.3.8}$$

且存在一些 ϑ 使上述严格不等式 $R(\vartheta, d_1) < R(\vartheta, d_2)$ 成立, 则称决策函数 $d_1(\boldsymbol{X})$ 一致优于 $d_2(\boldsymbol{X})$. 上式中 $\forall \vartheta \in \Theta$ 表示对一切 $\vartheta \in \Theta$ 不等式成立.

给定损失函数 $L(\vartheta, d)$, 设 $D = \{d(\boldsymbol{X})\}$ 是一切定义在样本空间 Ω_X 上而取值于决策空间 $\mathcal{D}$ 上的决策函数的全体, 若存在一个决策函数 $d^*(\boldsymbol{X}) \in D$, 对任意 $d(\boldsymbol{X}) \in D$ 下式成立：

$$R(\vartheta, d^*) \leqslant R(\vartheta, d), \forall \vartheta \in \Theta, \tag{13.3.9}$$

则称决策函数 $d^*(\boldsymbol{X})$ 为 (该决策函数类 D 的) 一致最小风险决策函数, 或一致最优决策函数.

上述两个定义都是针对某个给定的损失函数而言的, 当损失函数改变了, 相应的结论也随之而变.

以一个简单例子来说明以上这些概念的运用. 设总体 X 服从正态分布 $N(\mu, \sigma^2)$, σ 为已知, 欲用总体 X 的样本 $\boldsymbol{X} = (X_1, \cdots, X_n)^T$ 来估计未知参数 μ. 选取损失函数为 $L(\mu, d) = (d-\mu)^2$, 则对 μ 的任一估计 $d(\boldsymbol{X})$, 风险函数为

$$R(\mu, d) = E_\mu \left[L(\mu, d)\right] = E_\mu \left[(d-\mu)^2\right].$$

若进一步要求 $d(\boldsymbol{X})$ 是 μ 的无偏估计, 即 $E_\mu \left[d(\boldsymbol{X})\right] = \mu$, 则风险函数是

$$R(\mu, d) = E_\mu \left[(d - E(d))^2\right] = Var_\mu \left[d(\boldsymbol{X})\right],$$

即风险函数是估计量 $d(\boldsymbol{X})$ 的方差.

$$\text{若取 } d(\boldsymbol{X}) = \bar{X}, \text{ 则 } R(\mu, d) = Var_\mu \left[\bar{X}\right] = \frac{\sigma^2}{n},$$
$$\text{若取 } d(\boldsymbol{X}) = X_1, \text{ 则 } R(\mu, d) = Var_\mu \left[X_1\right] = \sigma^2,$$

显然, 当 $n > 1$ 时, 后者的风险大于前者, 即估计量 $\bar{X}$ 优于 X_1.

(3) 贝叶斯风险

贝叶斯统计中将参数 ϑ 视为参数空间 Θ 中具有先验分布 $\pi(\vartheta)$ 的随机变量, 因此前面定义的风险函数

$$R(\vartheta, d) = E_\vartheta \left[L(\vartheta, d(\boldsymbol{X}))\right] = \int_{\Omega_X} L(\vartheta, d(\boldsymbol{X})) f(\boldsymbol{X}|\vartheta) \mathrm{d}\boldsymbol{X}$$

是 ϑ 的函数, 仍然是随机变量, 需要在参数空间 Θ 中对于 ϑ 作期望值运算：

$$B(d) \equiv E[R(\vartheta, d)] = \int_\Theta R(\vartheta, d) \pi(\vartheta) \mathrm{d}\vartheta, \tag{13.3.10}$$

$B(d)$ 称为决策函数 d 在给定先验分布 $\pi(\vartheta)$ 下的贝叶斯风险, 简称 d 的贝叶斯风险.

当总体 X 和 ϑ 都是连续随机变量时, 上式可写为

$$\begin{aligned} B(d) &= \int_{\Theta} R(\vartheta, d)\pi(\vartheta)\mathrm{d}\vartheta \\ &= \int_{\Theta}\int_{\Omega_X} L(\vartheta, d(\boldsymbol{X}))f(\boldsymbol{X}|\vartheta)\pi(\vartheta)\mathrm{d}\boldsymbol{X}\mathrm{d}\vartheta \\ &= \int_{\Omega_X}\int_{\Theta} L(\vartheta, d(\boldsymbol{X}))q(\boldsymbol{X})h(\vartheta|\boldsymbol{X})\mathrm{d}\boldsymbol{X}\boldsymbol{d}\vartheta \\ &= \int_{\Omega_X} q(\boldsymbol{X})\left\{\int_{\Theta} L(\vartheta, d(X))h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta\right\}\mathrm{d}\boldsymbol{X}. \end{aligned} \tag{13.3.11}$$

当总体 X 和 ϑ 都是离散随机变量时, 则有

$$B(d) = \sum_{X} q(\boldsymbol{X})\left\{\sum_{\vartheta} L[\vartheta, d(\boldsymbol{X})]h(\vartheta|\boldsymbol{X})\right\}. \tag{13.3.12}$$

由此可见, 贝叶斯风险可视为随机损失函数求两次期望运算得到的, 第一次先对 ϑ 的后验分布求期望, 第二次对样本 $\boldsymbol{X}$ 的边沿分布求期望. 这时 $B(d)$ 已不依赖于参数 ϑ 而仅依赖于决策函数 $d(\boldsymbol{X})$.

13.3.2 贝叶斯参数点估计

从贝叶斯统计的观点来看, 对于参数 ϑ 的点估计问题, 若能根据总体 X 的随机子样 $\boldsymbol{X} = (X_1, \cdots, X_n)^T$ 找到 ϑ 的一个估计量 (一个决策)$d^*(\boldsymbol{X})$, 使得其贝叶斯风险 $B(d^*)$ 在所有的估计量的贝叶斯风险 $B(d) = \int_{\Theta} R(\vartheta, d)\pi(\vartheta)\mathrm{d}\vartheta$ 中达到极小值, 这个估计量 $d^*(\boldsymbol{X})$ 就是最佳的估计.

于是有下述定义: 设总体 X 的概率密度 $f(x, \vartheta)$ 中参数 ϑ 为随机变量, $\pi(\vartheta)$ 为 ϑ 的先验分布, $\boldsymbol{X} = (X_1, \cdots, X_n)^T$ 是 X 的随机子样. 若在决策空间 $\mathcal{D}$ 中存在一个决策函数 $d^*(\boldsymbol{X})$, 使得对决策空间 $\mathcal{D}$ 中任一决策函数 $d(\boldsymbol{X})$ 均有

$$B(d^*(\boldsymbol{X})) = \underset{d(X)}{Min}\, B(d(\boldsymbol{X})), \quad \forall d \in \mathcal{D}, \tag{13.3.13}$$

则称 $d^*(\boldsymbol{X})$ 是参数 ϑ 关于先验分布 $\pi(\vartheta)$ 的贝叶斯估计量, 或简称为**贝叶斯估计量**.

由定义可见, 贝叶斯估计量 $d^*(\mathbf{X})$ 是使贝叶斯风险 $B(d)$ 达到极小的决策函数.

应该注意, 贝叶斯估计量是依赖于先验分布 $\pi(\vartheta)$ 的, 即对于不同的 $\pi(\vartheta)$, ϑ 的贝叶斯估计量是不同的.

(1) 后验期望估计

若选定 ϑ 的先验分布 $\pi(\vartheta)$ 和平方差损失函数

$$L(\vartheta, d) = (\vartheta - d)^2,$$

则 ϑ 的贝叶斯估计量为

$$d(\boldsymbol{X}) = E(\vartheta|\boldsymbol{X}) = \int_{\Theta} \vartheta h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta \equiv \hat{\vartheta}_E. \tag{13.3.14}$$

其中 $\hat{\vartheta}_E$ 正是参数 ϑ 关于后验分布 $h(\vartheta|\boldsymbol{X})$ 的期望, 因此贝叶斯估计 $\hat{\vartheta}_E$ 称为参数 ϑ 的**后验期望估计**. 该式的证明参见文献 [92].

由于求得 $\hat{\vartheta}_E$ 时选择的是平方差损失函数, 因此当使用 $\hat{\vartheta}_E$ 作为参数 ϑ 的估计量时导致的方差损失最小. 后验期望估计是贝叶斯点估计方法中最重要的一种, 通常贝叶斯点估计指的是后验期望估计.

(2) 后验中位数估计

若选定 ϑ 的先验分布 $\pi(\vartheta)$ 和绝对差损失函数

$$L(\vartheta, d) = |\vartheta - d|,$$

则 ϑ 的贝叶斯估计量为后验分布 $h(\vartheta|X)$ 的中位数, 记为 $\hat{\vartheta}_m$, 称为**后验中位数估计**. 它的证明亦见文献 [92]. 由于求得 $\hat{\vartheta}_m$ 时选择的是绝对差损失函数, 因此当使用 $\hat{\vartheta}_m$ 作为参数 ϑ 的估计量时导致的绝对差损失最小.

(3) 后验最可几估计

若选定 ϑ 的先验分布 $\pi(\vartheta)$ 和 0, 1 损失函数

$$L(\vartheta, d) = \begin{matrix} 1, & d \neq \vartheta, \\ 0, & d = \vartheta. \end{matrix}$$

可以证明[93], ϑ 的贝叶斯估计量 (记为 $\hat{\vartheta}_p$) 对应于后验密度 $h(\vartheta|\boldsymbol{X})$ 最可几值 (即最大值) 处的 ϑ 值, 称为**后验最可几估计**, 即

$$h(\hat{\vartheta}_p|\boldsymbol{X}) = \underset{\vartheta \in \Theta}{Max}\, h(\vartheta|\boldsymbol{X}). \tag{13.3.15}$$

当采用先验分布的贝叶斯假设 $\pi(\vartheta) \propto 1$, 则由式 (13.2.9) 知

$$h(\vartheta|\boldsymbol{X}) \propto L(\boldsymbol{X}|\vartheta).$$

这时贝叶斯后验最可几估计与经典方法的极大似然估计相等.

由于 $\hat{\vartheta}_p$ 对应于后验密度 $h(\vartheta|\boldsymbol{X})$ 最大值处的 ϑ 值, 为求得 $\hat{\vartheta}_p$, 必须求 $h(\vartheta|\boldsymbol{X})$ 的最大值, 这可由求 $h(\vartheta|\boldsymbol{X})$ 对于 ϑ 的导数并令其等于 0 得到. 由于 $\ln h(\vartheta|\boldsymbol{X})$ 与

$h(\vartheta|\boldsymbol{X})$ 有相同的极值位置, 通常令 $\ln h(\vartheta|\boldsymbol{X})$ 对于 ϑ 的导数等于 0 计算更为方便. 例如欲求正态分布 $N(\mu_0,\vartheta),\vartheta=\sigma^2$ 的 $\hat{\vartheta}_p$, 由表 13.2 知当选取共轭先验分布 $I\Gamma(a,b)$ 时, 后验分布为

$$h(\vartheta|\boldsymbol{X})=I\Gamma\left(a+n/2,b+\sum_{i=1}^{n}(X_i-\mu_0)^2\bigg/2\right),$$

这时有

$$\ln h(\vartheta|\boldsymbol{X})=-\left(a+\frac{n}{2}+1\right)\ln\vartheta-\frac{1}{\vartheta}\left[b+\frac{1}{2}\sum_{i=1}^{n}(X_i-\mu_0)^2\right].$$

令 $\ln h(\vartheta|\boldsymbol{X})$ 对于 ϑ 的导数等于 0 求得后验最可几估计:

$$\hat{\vartheta}_p=\frac{2b+\sum\limits_{i=1}^{n}(X_i-\mu_0)^2}{2a+n+2}.$$

如果选取贝叶斯先验分布 $\pi(\vartheta)\propto 1$, 可将其形式地表示为 $\pi(\vartheta)\propto 1=I\Gamma(a,b)=I\Gamma(-1,0)$ 代入上式, 得到的是极大似然估计:

$$\hat{\vartheta}_{ML}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2.$$

利用类似的方法可求得表 13.2 所列的若干种常见概率分布参数的贝叶斯后验最可几估计 $\hat{\vartheta}_p$ 及极大似然估计 (选取贝叶斯先验分布), 列于表 13.3.

根据前面的讨论我们知道, 先验分布的贝叶斯假设仅对位置参数是恰当的, 因此对于位置参数在无信息条件下, 极大似然估计是合理的估计. 而对尺度参数族, 尺度参数 ϑ 在无信息条件下应取 $\pi(\vartheta)\propto 1/\vartheta$, 此时

$$h(\vartheta|\boldsymbol{X})\propto\frac{1}{\vartheta}L(\boldsymbol{X}|\vartheta),$$

后验最可几估计与极大似然估计不一致.

由以上讨论可知, 对于参数 ϑ 的点估计而言, 后验期望估计 $\hat{\vartheta}_E$ 是平均平方差损失最小的估计量, 后验中位数估计 $\hat{\vartheta}_m$ 是平均绝对差损失最小的估计量, 后验最可几估计 $\hat{\vartheta}_p$ 是平均 0, 1 损失最小的估计量. 一般情形下, 这三个估计值是不等的; 但当后验分布 $h(\vartheta|\boldsymbol{X})$ 为对称时, 这三个估计值相等.

我们注意到, 后验期望值、后验中位数和后验最可几值都是后验分布的特定的位置参数. 从这一现象来看, 贝叶斯统计推断问题中后验分布的作用与经典统计中似然函数的作用类似, 经典统计中参数 ϑ 的点估计正是对应于似然函数极大处的参数值. 两者的差别在于贝叶斯后验分布中包含了先验分布的信息, 这一点是经典统计中点估计方法不具备的.

表 13.3 常见概率分布参数的贝叶斯后验最可几估计 $\hat{\vartheta}_p$ 及极大似然估计 $\hat{\vartheta}_{ML}$

总体分布		先验分布	$\hat{\vartheta}_p$	$\hat{\vartheta}_{ML}$
二项分布 $\binom{n}{X}\vartheta^X(1-\vartheta)^{n-X}$		$\beta(a,b)$	$\frac{a+X-1}{a+b+n-2}$	$\frac{X}{n}$
几何分布 $\vartheta(1-\vartheta)^{X-1}$		$\beta(a,b)$	$\frac{a+n-1}{a+b+n\bar{X}-2}$	$\bar{X}^{-1}$
泊松分布 $\frac{1}{X!}\vartheta^X e^{-\vartheta}$		$\Gamma(a,b)$	$\frac{a+n\bar{X}-1}{b+n}$	$\bar{X}$
指数分布	$\vartheta e^{-\vartheta X}$	$\Gamma(a,b)$	$\frac{a+n-1}{b+n\bar{X}}$	$\bar{X}^{-1}$
	$\vartheta^{-1}e^{-X/\vartheta}$	$I\Gamma(a,b)$	$\frac{b+n\bar{X}}{a+n+1}$	$\bar{X}$
正态分布	$N(\vartheta,\sigma^2)$	$N(\mu,\tau^2)$	$\frac{\frac{nX}{\sigma^2}+\frac{\mu}{\tau^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}$	$\bar{X}$
	$N(\mu_0,\vartheta)$ $\vartheta=\sigma^2$	$I\Gamma(a,b)$	$\frac{2b+\sum_{i=1}^{n}(X_i-\mu_0)^2}{2a+n+2}$	$\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_0)^2$

表中 ϑ 为待估计参数, X 为测量值, $\bar{X}$ 和 n 为样本平均和样本容量, a,b,μ,τ 为超参数.

在后验期望估计、后验中位数估计和后验最可几估计之中, 究竟采用哪一种估计量要视问题的具体要求来确定; 但从估计的误差角度而言, 后验期望估计 $\hat{\vartheta}_E$ 性质最优. 下面我们来讨论这一问题.

(4) 贝叶斯估计的误差

设 $\hat{\vartheta}$ 是 ϑ 的一个贝叶斯估计, 在给定样本后, $\hat{\vartheta}$ 是一个数, 而随机变量 ϑ 按后验分布取值. 所以评定一个贝叶斯估计的误差的最好和方便的方法是用 ϑ 对 $\hat{\vartheta}$ 的后验均方差来度量. 于是有以下定义:

设参数 ϑ 的后验分布为 $h(\vartheta|\boldsymbol{X})$, 贝叶斯估计为 $\hat{\vartheta}$, 则 $(\hat{\vartheta}-\vartheta)^2$ 的后验期望

$$MSE(\hat{\vartheta}|\boldsymbol{X}) \equiv \int_{\Theta}(\hat{\vartheta}-\vartheta)^2 h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta \tag{13.3.16}$$

称为贝叶斯估计 $\hat{\vartheta}$ 的**后验均方差**. 后验均方差越小, 贝叶斯估计 $\hat{\vartheta}$ 的误差就越小.

当利用贝叶斯后验期望估计 $\hat{\vartheta}_E$ 作为参数 ϑ 的估计, 有

$$MSE(\hat{\vartheta}_E|\boldsymbol{X}) \equiv \int_{\Theta}(\hat{\vartheta}_E-\vartheta)^2 h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta \equiv Var(\vartheta|\boldsymbol{X}), \tag{13.3.17}$$

$Var(\vartheta|\boldsymbol{X})$ 称为**后验方差**, 其平方根 $[Var(\vartheta|\boldsymbol{X})]^{1/2}$ 称为**后验标准差**.

后验均方差与后验方差有如下关系:

$$MSE(\hat{\vartheta}|\boldsymbol{X}) \equiv \int_{\Theta}(\hat{\vartheta}-\vartheta)^2 h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta = \int_{\Theta}[(\hat{\vartheta}-\hat{\vartheta}_E)+(\hat{\vartheta}_E-\vartheta)]^2 h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta,$$

注意到在给定样本后, $\hat{\vartheta}_E$ 和 $\hat{\vartheta}$ 是具体的数, 故上式等于

$$MSE(\hat{\vartheta}|\boldsymbol{X}) = (\hat{\vartheta}_E - \hat{\vartheta})^2 + Var(\vartheta|\boldsymbol{X}). \tag{13.3.18}$$

这表明, 当利用 $\hat{\vartheta}_E$ 作为参数 ϑ 的估计, 可使后验均方差达到最小. 这与前面已经提到的使用 $\hat{\vartheta}_E$ 作为参数 ϑ 的估计量时导致的方差损失最小的结论是一致的. 所以在实际中常常取后验期望估计 $\hat{\vartheta}_E$ 作为参数 ϑ 的贝叶斯估计.

利用式 (13.3.14) 和式 (13.3.17), 求得几种常见概率分布参数的贝叶斯后验期望估计 $\hat{\vartheta}_E$ 及其后验标准差, 列于表 13.4. 其中用到了 $\beta(a,b)$, $\Gamma(a,b)$ 分布期望值和方差的性质 (见 4.9 节和 4.10 节), 以及 $I\Gamma(a,b)$ 分布期望值和方差的性质 (见式 (13.2.38)). 例如, 对于泊松分布参数, 当选取先验分布为 $\beta(a,b)$, 其后验分布由表 13.2 知为 $\Gamma(a+n\bar{X}, b+n)$. 由于 $\Gamma(a,b)$ 分布期望值和方差为 a/b 和 a/b^2, 立即可得后验分布的期望值 (即 $\hat{\vartheta}_E$) 及其后验标准差为

$$\frac{a+n\bar{X}}{b+n}, \quad \frac{\sqrt{a+n\bar{X}}}{b+n}.$$

通过类似的计算可求得表中其他分布的贝叶斯后验期望估计 $\hat{\vartheta}_E$ 及其后验标准差.

表 13.4 常见概率分布参数的贝叶斯后验期望估计 $\hat{\vartheta}_E$ 及其后验标准差

总体分布		先验分布	$\hat{\vartheta}_E$	后验标准差
二项分布 $\binom{n}{X}\vartheta^X(1-\vartheta)^{n-X}$		$\beta(a,b)$	$\dfrac{a+X}{a+b+n}$	$\left[\dfrac{(a+X)(b+n-X)}{(a+b+n)^2(a+b+n+1)}\right]^{1/2}$
几何分布 $\vartheta(1-\vartheta)^{X-1}$		$\beta(a,b)$	$\dfrac{a+n}{a+b+n\bar{X}}$	$\left[\dfrac{(a+n)(b+n\bar{X}-n)}{(a+b+n\bar{X})^2(a+b+n\bar{X}+1)}\right]^{1/2}$
泊松分布 $\dfrac{1}{X!}\vartheta^X e^{-\vartheta}$		$\Gamma(a,b)$	$\dfrac{a+n\bar{X}}{b+n}$	$\dfrac{\sqrt{a+n\bar{X}}}{b+n}$
指数分布	$\vartheta e^{-\vartheta X}$	$\Gamma(a,b)$	$\dfrac{a+n}{b+n\bar{X}}$	$\dfrac{\sqrt{a+n}}{b+n\bar{X}}$
	$\vartheta^{-1}e^{-X/\vartheta}$	$I\Gamma(a,b)$	$\dfrac{b+n\bar{X}}{a+n-1}$	$\dfrac{b+n\bar{X}}{(a+n-1)\sqrt{a+n-2}}$
正态分布	$N(\vartheta,\sigma^2)$	$N(\mu,\tau^2)$	$\dfrac{\dfrac{n\bar{X}}{\sigma^2}+\dfrac{\mu}{\tau^2}}{\dfrac{n}{\sigma^2}+\dfrac{1}{\tau^2}}$	$\dfrac{\sigma\tau}{\sqrt{n\tau^2+\sigma^2}}$
	$N(\mu_0,\vartheta)$ $\vartheta=\sigma^2$	$I\Gamma(a,b)$	$\dfrac{2b+\sum\limits_{i=1}^{n}(X_i-\mu_0)^2}{2a+n-2}$	$\dfrac{2b+\sum\limits_{i=1}^{n}(X_i-\mu_0)^2}{(2a+n-2)\sqrt{a+\dfrac{n}{2}-2}}$

表中, ϑ 为待估计参数, X 为测量值, $\bar{X}$ 和 n 为样本平均和样本容量,a,b,μ,τ 为超参数.

我们通过一个具体例子来比较贝叶斯后验期望估计与极大似然估计的差别.

例 13.13 探测效率的贝叶斯后验期望估计与极大似然估计

一个粒子射入探测器得到一次计数的概率称为探测效率. 设 n 个粒子射入某探测器, 给出 x 个计数信号. 假定我们此前对探测效率 ϑ 没有先验知识, 试求探测效率 ϑ 的极大似然估计 $\hat{\vartheta}_{ML}$ 和贝叶斯后验期望估计 $\hat{\vartheta}_E$, 并比较它们的误差.

解 n 个粒子射入探测器给出 x 个计数信号, x 服从参数 ϑ 的二项分布, 显然 ϑ 即是待估计的探测效率. 对 ϑ 没有先验知识意味着 ϑ 的值在区间 [0,1] 之间没有任何倾向性, 应当采用贝叶斯先验分布即 $\pi(\vartheta) = 1 = \beta(a,b) = \beta(1,1)$. 根据表 13.4 可知贝叶斯后验期望估计 $\hat{\vartheta}_E$ 及其后验标准差为

$$\hat{\vartheta}_E = \frac{a+x}{a+b+n} = \frac{1+x}{2+n},$$

$$\sigma_{\vartheta_E} = \left[\frac{(a+x)(b+n-x)}{(a+b+n)^2(a+b+n+1)}\right]^{1/2} = \left[\frac{(1+x)(1+n-x)}{(2+n)^2(3+n)}\right]^{1/2}.$$

而据表 13.3 可知极大似然估计为 $\hat{\vartheta}_{ML} = x/n$, 其标准差由第 4 章例 4.1 的讨论知为

$$\sigma_{\vartheta_{ML}} = \left[\frac{x}{n^2}\left(1-\frac{x}{n}\right)\right]^{1/2}.$$

两者对于 ϑ 及其误差的估计是不同的.

我们代入具体数据来体察两者的差异. 例如对于 n=10, x=2,8 以及 n=1000, x=20,80,200,800 这两组数据, 可得表 13.5:

表 13.5 探测效率的极大似然估计 $\hat{\vartheta}_{ML}$ 和贝叶斯后验期望估计 $\hat{\vartheta}_E$ 的比较

n	x	$\hat{\vartheta}_{ML}$	$\hat{\vartheta}_E$
10	2	0.20 ± 0.13	0.25 ± 0.12
	8	0.80 ± 0.13	0.75 ± 0.12
1000	20	0.0200 ± 0.0044	0.0210 ± 0.0045
	80	0.0800 ± 0.0086	0.0808 ± 0.0086
	200	0.2000 ± 0.0127	0.2006 ± 0.0127
	800	0.8000 ± 0.0127	0.7994 ± 0.0127

根据以上结果, 可有下述结论: ① $\hat{\vartheta}_{ML}$ 和 $\hat{\vartheta}_E$ 在误差范围内一致. ② 当 n 和/或 x 小时, 两者中心值差别明显; 当 n 和 x 大时, 两者中心值差别很小.

究竟 $\hat{\vartheta}_{ML}$ 和 $\hat{\vartheta}_E$ 哪个估计更加合理? 为此我们来讨论一种具体情况 $x = n$, 即 n 个粒子全部探测到的情形. 这时极大似然估计为

$$\hat{\vartheta}_{ML} = 1, \quad \sigma_{\hat{\vartheta}_{ML}} = 0.$$

而贝叶斯后验期望估计为

$$\hat{\vartheta}_E = 1 - \frac{1}{2+n}, \quad \sigma_{\hat{\vartheta}_E} = \frac{1}{n+2}\sqrt{\frac{n+1}{n+3}}.$$

显然, 极大似然估计具有明显的不合理性, 因为 n 个粒子全部探测到并不能保证 $n+1, n+2, \cdots$ 个粒子也能全部探测到, 即有限个粒子全部探测到不能保证探测效率为 1. 同时该估计量的误差为 0 也不合理. 与此不同, 在 n 个粒子全部探测到的情形下, 随着 n 的增大, $\hat{\vartheta}_E$ 趋近于 1, 而误差 σ_{ϑ_E} 趋近于 0, 这都是符合常理的. 因此在粒子探测效率及其误差的确定上, 贝叶斯后验期望估计 $\hat{\vartheta}_E$ 比极大似然估计 $\hat{\vartheta}_{ML}$ 要合理.

本例中参数 ϑ 代表探测效率, ϑ 是二项分布的参数, 因此以上讨论对于所有服从二项分布的情形都适用. 例如产品的合格率同样可视为二项分布的参数. 在探测效率的测量中, 通常 n 和 x 都很大, 因此利用 $\hat{\vartheta}_E$ 和 $\hat{\vartheta}_{ML}$ 作为估计没有什么差别; 而在产品合格率的检测中, 样本量 n 通常都不大, 由上述讨论可知, 利用贝叶斯后验期望估计 $\hat{\vartheta}_E$ 更为合理.

13.3.3　经验贝叶斯估计

贝叶斯估计的缺点是先验分布往往很难确定, 当先验分布未知或先验分布中包含未知超参数时就无法找到贝叶斯估计. 罗宾斯 (Robbins) 于 1955 年提出**经验贝叶斯方法**(empirical Bayes method), 在一定程度上克服了这一困难[91].

设总体 X 的分布函数为 $F(X|\vartheta)$, 其中 ϑ 为未知参数. 假定 ϑ 随时间而变化, 故某一时刻 ϑ 的取值是随机的. 假定过去已经获得总体 X 的 m 批样本 $\boldsymbol{X}_1 = \left\{\boldsymbol{X}_1^{(1)}, \cdots, \boldsymbol{X}_{n_1}^{(1)}\right\}, \cdots, \boldsymbol{X}_m = \left\{\boldsymbol{X}_1^{(m)}, \cdots, \boldsymbol{X}_{n_m}^{(m)}\right\}$, 那么这 m 批样本包含了过去 ϑ 值的信息. 设当前取得的样本值为 $\boldsymbol{X}_{m+1} = \left\{X_1^{(m+1)}, \cdots, X_{n_{m+1}}^{(m+1)}\right\}$, 现在的问题是怎样用历史样本和当前样本来估计 ϑ 的当前值.

经验贝叶斯方法解决这一问题的基本思想是: 利用历史样本 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_m$ 来估计 ϑ 的先验分布, 得到先验分布 $\pi(\vartheta)$ 的一个估计 $\hat{\pi}(\vartheta)$. 然后把 $\hat{\pi}(\vartheta)$ 作为 ϑ 的先验分布 $\pi(\vartheta)$, 与当前样本值 $\boldsymbol{X}_{m+1}$ 结合在一起用通常的方法构建后验分布, 从而求得 “经验贝叶斯估计”.

显然, 先验分布的估计 $\hat{\pi}(\vartheta)$ 与真正的先验分布 $\pi(\vartheta)$ 不一定完全相同, 因此经验贝叶斯方法得到的 “经验贝叶斯估计” 不一定是真正的贝叶斯估计. 同时, 对于同一批历史样本 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_m$, 用不同的估计方法会得到不同形式的 $\hat{\pi}(\vartheta)$, 因此会有不同的 “经验贝叶斯估计”.

下面介绍两种从实用观点看可行的经验贝叶斯估计方法.

(1) 第 1 种经验贝叶斯估计方法

适用第 1 种经验贝叶斯估计方法的前提是, 参数 ϑ 的先验分布 $\pi(\vartheta)$ 的形式为已知, 但其中包含未知的超参数 λ. 因此, ϑ 的先验分布可表示为 $\pi_\lambda(\vartheta)$. 用历史样本 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_m$ 来估计超参数 λ 得到其估计值 $\hat{\lambda} = \hat{\lambda}(\boldsymbol{X}_1, \cdots, \boldsymbol{X}_m)$, 从而得到先验分布的一个估计 $\pi_{\hat{\lambda}}(\vartheta)$. 将 $\pi_{\hat{\lambda}}(\vartheta)$ 作为先验分布, 结合当前样本 X_{m+1} 构建经验

贝叶斯估计.

例 13.14 探测效率的经验贝叶斯估计

k 个粒子穿过某探测器, 其中 x 个粒子被记录, x 服从参数 ϑ 的二项分布：

$$B(x;k,\vartheta)=\begin{pmatrix}k\\x\end{pmatrix}\vartheta^x(1-\vartheta)^{k-x},\quad x=0,1,\cdots,k.$$

其中参数 ϑ 为探测器的探测效率. 探测效率取决于探测器的工作状态, 例如气体探测器的气体成分、工作高压等等, 因此长时间内不同时刻的 ϑ 为一随机变量. 假定其先验分布为贝塔分布 $\beta(\vartheta;a,b)$, 其中超参数 a, b 未知. 假定历史样本为 $x_1,\cdots,x_n$, 当前样本为 x_{n+1}. 根据历史样本可求得其样本均值和样本方差, 作为贝塔分布均值和方差的估计：

$$\begin{cases}\bar{x}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}x_i=E(\beta)=\dfrac{a}{a+b},\\ S^2=\dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(x_i-\bar{x})^2=V(\beta)=\dfrac{ab}{(a+b)^2(a+b+1)}.\end{cases}\tag{13.3.19}$$

解此方程组即得超参数 a, b 的估计值 $\hat{a},\hat{b}$, 则先验分布的一个估计是 $\beta\left(\vartheta;\hat{a},\hat{b}\right)$. 由表 13.2 知参数 ϑ 的后验分布为 $\beta\left(\hat{a}+x_{n+1},\hat{b}+k-x_{n+1}\right)$, 由表 13.4 知参数 ϑ 的当前经验贝叶斯估计 (贝叶斯后验期望估计) 为

$$\hat{\vartheta}_{n+1}=\frac{\hat{a}+x_{n+1}}{\hat{a}+\hat{b}+k},\tag{13.3.20}$$

该估计的误差 (后验标准差) 为

$$\sigma_{n+1}=\left[\frac{(\hat{a}+x_{n+1})\left(\hat{b}+k-x_{n+1}\right)}{\left(\hat{a}+\hat{b}+k\right)^2\left(\hat{a}+\hat{b}+k+1\right)}\right]^{1/2}.\tag{13.3.21}$$

以上的结论对于总体服从参数 ϑ 的二项分布, 而先验分布取为 $\beta(a,b)$ 的问题均适用, 例如长时间的产品合格率问题可用同样的方法求得.

(2) 第 2 种经验贝叶斯估计方法

如果参数 ϑ 的先验分布的形式也未知, 则适用第 2 种经验贝叶斯估计方法. 根据参数 ϑ 在总体分布中的特征性质, 用通常的估计方法从历史样本 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m$ 出发求得对参数 ϑ 的 m 个估计 $\hat{\vartheta}_1=\hat{\vartheta}(\boldsymbol{X}_1),\cdots,\hat{\vartheta}_m=\hat{\vartheta}(\boldsymbol{X}_m)$. 将这 m 个估计视为参数 ϑ 的容量 m 的随机子样, 由它们可构造 ϑ 的子样分布函数或经验分布函数

$F_m^*(\vartheta)$(见式 (6.1.4)):

$$F_m^*(\vartheta)=\begin{cases}0, & \vartheta<\hat{\vartheta}_1,\\ k/n, & \hat{\vartheta}_k\leqslant\vartheta<\hat{\vartheta}_{k+1}, \quad k=1,\cdots,m-1,\\ 1, & \vartheta\geqslant\hat{\vartheta}_m.\end{cases}\tag{13.3.22}$$

用 $F_m^*(\vartheta)$ 作为先验分布的近似. $F_m^*(\vartheta)$ 只在 $\hat{\vartheta}_1,\cdots,\hat{\vartheta}_m$ 上有概率, 且是等概率的, 即相应的离散概率为

$$\pi\left(\hat{\vartheta}_i\right)=1/m,\quad i=1,\cdots,m.\tag{13.3.23}$$

设总体的概率密度为 $f(x|\vartheta)$, 当前样本 $\boldsymbol{X}_{m+1}$ 的似然函数为

$$L\left(\boldsymbol{X}_{m+1}|\vartheta\right)=\prod_{j=1}^{n_{m+1}}f\left(X_j|\vartheta\right),\tag{13.3.24}$$

其中 n_{m+1} 为当前样本 $\boldsymbol{X}_{m+1}$ 的容量. 于是 ϑ 与当前样本 $\boldsymbol{X}_{m+1}$ 的联合 (经验) 频率函数为

$$\pi(\vartheta)L\left(\boldsymbol{X}_{m+1}|\vartheta\right)=\frac{1}{m}\sum_{i=1}^{m}I_{\vartheta=\hat{\vartheta}_i}L\left(\boldsymbol{X}_{m+1}|\vartheta\right),\tag{13.3.25}$$

其中 $I_{\vartheta=\hat{\vartheta}_i}$ 为示性函数, 即

$$I_{\vartheta=\hat{\vartheta}_i}=\begin{cases}1, & \vartheta=\hat{\vartheta}_i,\\ 0, & \vartheta\neq\hat{\vartheta}_i.\end{cases}\tag{13.3.26}$$

于是 ϑ 的后验频率函数为

$$\hat{h}(\vartheta)=\frac{\displaystyle\sum_{i=1}^{m}I_{\vartheta=\hat{\vartheta}_i}L\left(\boldsymbol{X}_{m+1}|\vartheta\right)}{\displaystyle\sum_{i=1}^{m}L\left(\boldsymbol{X}_{m+1}|\hat{\vartheta}_i\right)}.\tag{13.3.27}$$

这时贝叶斯估计为

$$\hat{\vartheta}_E=\int\vartheta\hat{h}(\vartheta)\mathrm{d}\vartheta=\frac{\displaystyle\sum_{i=1}^{m}\hat{\vartheta}_iL\left(\boldsymbol{X}_{m+1}|\hat{\vartheta}_i\right)}{\displaystyle\sum_{i=1}^{m}L\left(\boldsymbol{X}_{m+1}|\hat{\vartheta}_i\right)}.\tag{13.3.28}$$

举一个数值例子, 检验产品合格率, 历史上检验了 6 次, 10 个产品中合格产品数为 x= 8, 7, 8, 6, 9, 10, 本次合格产品数为 9, 产品合格率 ϑ 的先验分布未知, 要求本次合格率的经验贝叶斯估计.

显然 x 服从参数 ϑ 的二项分布, 即总体分布为 $C_k^x\vartheta^x(1-\vartheta)^{k-x}$, $k=10$, 用通常的 (极大似然) 估计法从历史样本 $x_1,\cdots,x_6$ 求得对参数 ϑ 的 $m=6$ 个估计值 $\left\{\hat{\vartheta}_1,\cdots,\hat{\vartheta}_6\right\}=\{0.8,0.7,0.8,0.6,0.9,1.0\}$, 本次测量的似然函数为

$$L(x_7|\vartheta)=f(x_7|\vartheta)=C_{10}^9\vartheta^9(1-\vartheta)=10\vartheta^9(1-\vartheta)$$

于是据式 (13.3.28), 本次合格率的经验贝叶斯估计为

$$\hat{\vartheta}_E=\frac{\sum\limits_{i=1}^{6}\hat{\vartheta}_i^{10}\left(1-\hat{\vartheta}_i\right)}{\sum\limits_{i=1}^{6}\hat{\vartheta}_i^{9}\left(1-\hat{\vartheta}_i\right)}=\frac{0.0887}{0.1086}=0.817.$$

13.3.4　贝叶斯参数区间估计

经典统计中求未知参数 ϑ 的区间估计方法受到相当多的批评. 一方面, 它需要构造一个枢轴量, 它是未知参数 ϑ 的函数, 但它的分布却要与 ϑ 无关, 对于一般的总体分布而言, 构建枢轴量相当困难. 另一方面, 经典统计中置信区间含义的理解存在困难. 若未知参数 ϑ 的置信水平 γ 的置信区间为 $(\vartheta_L,\vartheta_U)$, 这时可写 $P(\vartheta_L\leqslant\vartheta\leqslant\vartheta_U)=\gamma$. 但是在经典统计中, ϑ 为一未知常数, 它或者在区间 $(\vartheta_L,\vartheta_U)$ 内, 或者不在区间 $(\vartheta_L,\vartheta_U)$ 内, ϑ 落在区间 $(\vartheta_L,\vartheta_U)$ 内的概率不可能是 γ. 因此经典统计只能从频率的角度给予解释, 即对总体做 N 次抽样, 利用经典统计方法求得 N 个置信水平 γ 的置信区间, 当 N 充分大, 这 N 个区间包含 ϑ 真值的比率为 γ.

在贝叶斯统计中, 未知参数 ϑ 理解为随机变量, 它具有后验分布 $h(\vartheta|\boldsymbol{X})$, $\boldsymbol{X}$ 是样本观测值, 因此容易确定未知参数 ϑ 落入某一区间的后验概率. 因此, 贝叶斯区间估计的解释是非常自然的, 其区间估计方法比经典方法容易和直接.

对于给定的**信度概率**(credible probability) 或**信度水平**(credible level–CL)γ, 满足下式的区间 $[\vartheta_L,\vartheta_U]$ 即为贝叶斯信度区间 (credible interval):

$$P(\vartheta_L\leqslant\vartheta\leqslant\vartheta_U|\boldsymbol{X})=\gamma,\quad \vartheta\text{为连续量} \tag{13.3.29}$$

$$P(\vartheta_L\leqslant\vartheta\leqslant\vartheta_U|\boldsymbol{X})\geqslant\gamma,\quad \vartheta\text{为离散量} \tag{13.3.30}$$

因此, 贝叶斯信度区间的概率含义亦十分清晰, 没有任何含糊之处.

对于给定的信度水平 γ, 式 (13.3.29) 和式 (13.3.30) 定义的信度区间有无穷多个. 因此, 对于给定的信度水平 γ, 如何求得贝叶斯意义下的最优区间仍然是一个问题.

已知参数 ϑ 具有后验分布 $h(\vartheta|\boldsymbol{X})$, $\boldsymbol{X}$ 是样本观测值, 对于给定的信度水平 γ, 若存在区间 I, 满足下列条件:

$$P[\vartheta \in I|\boldsymbol{X}] = \int_I h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta = \gamma, \quad \vartheta\text{为连续量}, \tag{13.3.31}$$

$$P[\vartheta \in I|\boldsymbol{X}] = \sum_{\vartheta_i \in I} h(\vartheta_i|\boldsymbol{X}) \geqslant \gamma, \quad \vartheta\text{为离散量}. \tag{13.3.32}$$

且对任意 $\vartheta_1 \in I$, $\vartheta_2 \notin I$, 总有

$$h(\vartheta_1|\boldsymbol{X}) \geqslant h(\vartheta_2|\boldsymbol{X}), \tag{13.3.33}$$

则称 I 是参数 ϑ 的信度水平 γ 的**最大后验密度 (HPD) 区间**估计.

式 (13.3.33) 表示区间 I 内集中了后验密度取值尽可能大的点, 因此 ϑ 的 HPD 区间一定是相同信度水平 γ 下长度最短的区间. 一般认为 HPD 区间是贝叶斯意义下的最优区间估计.

对于给定的信度水平 γ, 下式决定参数 ϑ 的**贝叶斯信度区间上限**ϑ_{UP}:

$$P[\vartheta \leqslant \vartheta_{UP}|\boldsymbol{X}] = \int_{\vartheta \leqslant \vartheta_{UP}} h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta = \gamma, \quad \vartheta\text{为连续量}, \tag{13.3.34}$$

$$P[\vartheta \leqslant \vartheta_{UP}|X] = \sum_{\vartheta_i \leqslant \vartheta_{UP}} h(\vartheta_i|X) \geqslant \gamma, \quad \vartheta\text{为离散量}. \tag{13.3.35}$$

也可以构建**中心信度区间**$(\boldsymbol{\vartheta_L}, \boldsymbol{\vartheta_U})$:

$$P[\vartheta \leqslant \vartheta_L|\boldsymbol{X}] = P[\vartheta \geqslant \vartheta_U|\boldsymbol{X}] = \frac{1-\gamma}{2}, \quad \vartheta\text{为连续量}, \tag{13.3.36}$$

$$P[\vartheta \leqslant \vartheta_L|\boldsymbol{X}] = P[\vartheta \geqslant \vartheta_U|\boldsymbol{X}] \leqslant \frac{1-\gamma}{2}, \quad \vartheta\text{为离散量}. \tag{13.3.37}$$

当后验密度 $h(\vartheta|\boldsymbol{X})$ 为单峰、对称函数时, 中心信度区间即是最大后验密度 (HPD) 区间.

经典统计学中枢轴量的构建和寻找有时是相当困难的, 但寻求参数 ϑ 的贝叶斯信度区间只要利用 ϑ 的后验分布, 而不需要再去寻求其他的分布. 两相比较, 贝叶斯信度区间的寻找较为简单. 但是需要强调指出, 贝叶斯信度区间的确定依赖于后验分布, 从而也依赖于先验分布. 换言之, 如果选择不同的先验分布, 对于同样的信度概率, 其贝叶斯信度区间是不同的.

另一点需要强调的是经典置信区间置信概率 γ(或 CL) 与贝叶斯信度区间的信度概率 γ(或 CL) 的含义是不相同的, 当两者取相同值时, 置信区间和信度区间的长度可以是不同的.

例 13.15 正态总体均值的信度区间

设 $\boldsymbol{X}=(X_1,\cdots,X_n)^T$ 为正态总体 $N(\mu,\sigma^2)$ 的随机子样, σ^2 已知, μ 为未知参数. 求 μ 的信度概率 $\gamma=1-\alpha$ 的信度区间.

选取贝叶斯假设, 即 $\pi(\mu)\propto 1$, 此时如例 13.1 中所述, 参数 μ 的后验分布就是正态分布 $h\left(\mu|\bar{X}\right)=N\left(\bar{X},\sigma^2/n\right)$, 于是由后验分布知 $(\mu-\bar{X})\Big/\sqrt{\sigma^2/n}\sim N(0,1)$. 故有

$$P\left\{\left|\frac{\mu-\bar{X}}{\sigma/\sqrt{n}}\right|\leqslant z_{\alpha/2}\right\}=1-\alpha=\gamma,$$

其中 $z_{\alpha/2}$ 是标准正态分布的双侧 α 分位数. 由此得到 μ 的信度概率 $\gamma=1-\alpha$ 的信度区间为

$$\left[\bar{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\bar{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

此区间与经典方法得到的置信区间一致, 它实质上反映了没有任何先验信息可利用, 只能靠样本提供的信息来估计.

例 13.16　本底分布上叠加小信号情况下信号事例数的确定

第 11 章讨论了经典统计架构下粒子物理实验中小信号测量的区间估计问题. 现在讨论怎样利用贝叶斯统计方法来处理这类问题. 借用第 11 章中北京谱仪国际合作组利用 $\mathrm{e}^+\mathrm{e}^-$ 对撞机测量质心能量 $E_{\mathrm{cm}}=3650$, $3686, 3773\mathrm{MeV}$ 处 $\mathrm{e}^+\mathrm{e}^-\to\rho\eta'$ 的反应截面的测量数据图 11.1, 在 $E_{\mathrm{cm}}=3686\mathrm{MeV}$ 处不变质量 $M_{\eta\pi^+\pi^-}$ 的分布如图 11.1(b), 其中可以看到用多项式描述的本底分布 (用 $f_b(m)$ 表示其概率密度), 以及 958MeV 处的 η' 粒子信号. 显而易见, 观测到的事例主要来自于本底的贡献而信号事例数很少, 这种大本底上叠加小信号的情形在粒子物理实验测量中具有典型性. 要求利用贝叶斯方法求信号事例数 s 在信度概率 $\gamma=1-\alpha$ 下的信度区间.

首先确定一个基本包含全部信号事例的 “信号区间”, 例如现在的情形下可以是 [870,1060] MeV 区间. 可以认为该区间内本底事例数为期望值 b 的泊松变量, 期望值 b 可以用图 11.1(b) 中本底分布的拟合函数 $f_b(m)$ 求得. 该区间内信号事例数为期望值 s 的泊松变量, 期望值 s 为待估计参数. 这时, 该区间内观测事例数 n 为期望值 $s+b$ 的泊松变量, 即观测到 n 个事例的概率为

$$p(n|s)_b=\mathrm{e}^{-(s+b)}\frac{(s+b)^n}{n!}.\tag{13.3.38}$$

贝叶斯统计中, 参数估计基于后验概率：

$$h(s|n)=\frac{q(n|s)_b\pi(s)}{\displaystyle\int_0^\infty q(n|s)_b\pi(s)\mathrm{d}s},\tag{13.3.39}$$

其中 $q(n|s)_b$ 为给定 s 条件下观测到 n 个事例的条件概率, 亦即式 (13.2.9) 中的似然函数 $L\left(\boldsymbol{X}|\vartheta\right)=L\left(n|s\right)$, 在目前情形下即为式 (13.3.38) 所示的泊松变量

$p(n|s)_b$；$\pi(s)$ 是无信息先验概率：

$$\pi(s) \propto \frac{1}{(s+b)^m}, \quad s \geqslant 0, \quad b \geqslant 0 \quad 0 \leqslant m \leqslant 1, \tag{13.3.40}$$

m=0 对应于贝叶斯先验分布, m=0.5 对应于 $1/\sqrt{s+b}$ 先验分布 (Jeffreys 先验分布), m=1 对应于 $1/(s+b)$ 先验分布. 当选择不同的 m 值时得到不同的信度区间. 文献 [95] 对不同的 m 值得到不同的信度区间进行了讨论, 并论证了 $1/\sqrt{s+b}$ 先验分布给出了泊松变量信度区间的信度水平与其对应的经典置信水平的平均值相近. 图 13.1 显示了本底 b=0 的泊松变量, CL=68%和 90%两种情形下两者的对比.

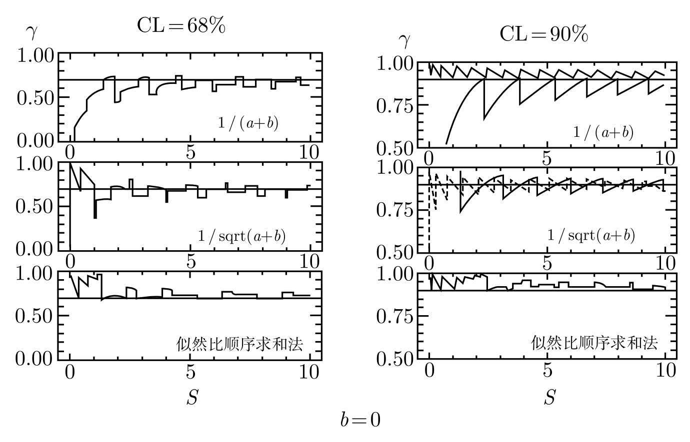


图 13.1　3 种先验分布构建的、信度水平为 CL 的贝叶斯信度区间所对应的经典置信概率 (γ).

式 (13.3.38) 所示的泊松变量中 b=0, 图中似然顺序求和法对应于先验分布 $\pi=(s)\propto 1$

于是在先验分布为 $\pi(s) \propto \dfrac{1}{(s+b)^m}$ 的情形下, 有

$$h(s|n) = \frac{(s+b)^{n-m}\mathrm{e}^{-(s+b)}}{\Gamma(n-m+1,b)}, \tag{13.3.41}$$

其中 $\Gamma(x,b)$ 是不完全 gamma 函数：

$$\Gamma(x,b) = \int_b^{\infty} s^{x-1}\mathrm{e}^{-s}\mathrm{d}s, \quad x>0, b>0. \tag{13.3.42}$$

因此可求得信度水平 $CL=1-\alpha$ 的中心区间 $[S_L, S_U]$：

$$\int_0^{S_L} h(s|n)\mathrm{d}s = \frac{\alpha}{2} = \int_{S_U}^{\infty} h(s|n)\mathrm{d}s, \tag{13.3.43}$$

以及上限 S_{UP}：

$$1-\alpha = \int_0^{S_{\mathrm{UP}}} h(s|n)\mathrm{d}s. \tag{13.3.44}$$

将式 (13.3.41) 代入上式求得

$$\alpha = \frac{\Gamma(n-m+1, S_{\text{UP}}+b)}{\Gamma(n-m+1, b)}. \tag{13.3.45}$$

当先验概率为均匀分布 ($m=0$), 上式变为

$$\alpha = \mathrm{e}^{-S_{\text{UP}}} \cdot \sum_{i=0}^{n} \frac{(S_{\text{UP}}+b)^i}{i!} \bigg/ \sum_{i=0}^{n} \frac{b^i}{i!}. \tag{13.3.46}$$

该式需数值求解, 步骤如下：式 (13.3.46) 可改写为

$$S_{\text{UP}} = \ln I(S_{\text{UP}}) - \ln\alpha,$$

其中,

$$I(S_{\text{UP}}) = \sum_{i=0}^{n} \frac{(S_{\text{UP}}+b)^i}{i!} \bigg/ \sum_{i=0}^{n} \frac{b^i}{i!}.$$

通过对目标函数

$$F = |S_{\text{UP}} - \ln I + \ln\alpha| \text{ 或 } \quad F = (S_{\text{UP}} - \ln I + \ln\alpha)^2$$

求极小可得到上限 S_{UP} 的解.

最大后验密度 (HPD) 信度区间 R 则由下式求得

$$1-\alpha = \int_R h(s|n)\mathrm{d}s, \tag{13.3.47}$$

其中对任意 $s_1 \in R$ 和 $s_2 \notin R$, 下述不等式成立

$$h(s_1|n) \geqslant h(s_2|n). \tag{13.3.48}$$

如前所述, 给定 CL 下, HPD 区间 $[S_L, S_U]$ 长度最短, 而且根据测量数据会自动给出信度区间或上限, 而无需人为地加以规定.

当需要考虑本底期望值和信号探测效率测量值的系统误差时, 本底期望值和信号探测效率不是常数, 而是随机变量, 服从某种分布. 这时式 (13.3.39) 的后验概率中的 $q(n|s)_b$ 不能用简单的泊松分布表示而需要修改. 如果只考虑本底的系统误差, 本底期望值 (改为 b' 表示) 的 pdf 表示为 $f'_b(b, \sigma_b)$, 其均值为 b 而标准差为 σ_b, 则条件概率密度 $q(n|s)_b$ 须修改为

$$q(n|s)_b = \int_0^\infty p(n|s)_{b'} \cdot f_{b'}(b, \sigma_b)\mathrm{d}b', \tag{13.3.49}$$

式中 $p(n|s)_{b'}$ 与式 (13.3.38) 相同, 但 b 用 b' 代替.

如果同时考虑信号效率和本底的系统误差, 且两者独立, 相对信号效率 ε 的 *pdf* 表示为 $f_\varepsilon(1,\sigma_\varepsilon)$, 其均值为 1, 标准差为 σ_ε, 则条件概率密度 $q(n|s)_b$ 须修改为

$$q(n|s)_b = \int_0^\infty \int_0^\infty p(n|s\varepsilon)_{b'} f_{b'}(b,\sigma_b) f_\varepsilon(1,\sigma_\varepsilon)\mathrm{d}b'\mathrm{d}\varepsilon, \tag{13.3.50}$$

式中 $p(n|s\varepsilon)'_b$ 与式 (13.3.38) 相同, 但 b 用 b' 代替, s 用 $s\varepsilon$ 代替.

用式 (13.3.49) 或式 (13.3.50) 代入式 (13.3.39) 即求得后验概率 $h(s|n)$, 其余步骤相同.

文献 [97],[98] 详尽地讨论了泊松观测量情形下贝叶斯信度区间的构建问题, 并编制了计算贝叶斯 HPD 信度区间和上限的 FORTRAN 程序包 BPOCI (Bayesian POissonian Credible Interval) 供使用 [99]. 图 13.2 是观测事例数 n=8 情形下用 BPOCI 程序包算得的后验分布 $h(s|n)$ 与 s 的函数关系, 其他条件见图中的说明.

由此求得的 HPD 信度区间 (CL=0.9) 为 [2.07, 11.77].

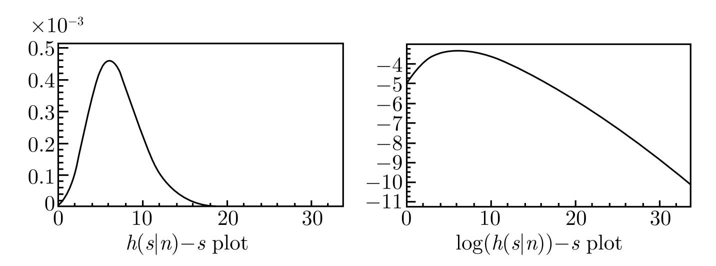


图 13.2 BPOCI 程序包算得的后验分布 $h(s|n)$ 与 s 的函数关系, n=8,
左图为线性坐标, 右图为对数坐标

利用均匀分布先验密度 (即 $m=1$), 本底期望值考虑为高斯分布, 均值 b= 2, σ_b/b=0.3

书末的附表 23 给出了先验分布为均匀分布情形下, 信号区内信号泊松事例数期望值 s 的 HPD 信度区间.

给定信度概率下, 贝叶斯 HPD 区间长度最短, 而且根据测量数据会自动给出信度区间或上限, 而无需人为地加以规定；同时, 式 (13.3.31) 和式 (13.3.32) 保证了不存在信度概率不足的问题, 也不存在空集的困难, 因此贝叶斯 HPD 区间克服了经典统计 Neyman 方法在小信号区间估计中的所有困难. 与 11 章中讨论的似然比顺序求和方法 (FC 方法) 相比, 对比该方法给出的书末附表 10 和贝叶斯 HPD 区间的附表 23, 可以发现, 对于泊松观测量中 n 接近或小于 b 的情形, 贝叶斯 HPD 区间对于本底期望值 b 的依赖要远小于 FC 方法确定的置信区间对于 b 的依赖.

13.3.5 贝叶斯假设检验

在贝叶斯统计中, 假设检验问题比较容易处理. 设总体 X 的概率密度 $f(X;\vartheta)$ 函数形式为已知, 但其中包含未知参数 ϑ. 要求从总体的子样测量值 $\boldsymbol{X} = (X_1,$

$X_2, \cdots, X_n)^T$ 来决定究竟应当接受零假设 $H_0 : \vartheta \in \Theta_0$ 还是接受备择假设 $H_1 : \vartheta \in \Theta_1$. 其中 Θ_0, Θ_1 为零假设 H_0 和备择假设 H_1 为真时的参数空间, 且 $\Theta_0 \cap \Theta_1 = \varnothing$, 即 Θ_0, Θ_1 无公共部分.

假定我们选定了先验分布为 $\pi(\vartheta)$, 并根据子样测量值 $\boldsymbol{X}$ 和式 (13.2.9) 得到后验分布 $h(\vartheta|\boldsymbol{X})$. 记 α_0 和 α_1 为零假设 H_0 和备择假设 H_1 相应的后验概率：

$$\alpha_0 = \alpha_0(\boldsymbol{X}) = P(\vartheta \in \Theta_0|\boldsymbol{X}) = \int_{\Theta_0} h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta, \tag{13.3.51}$$

$$\alpha_1 = \alpha_1(\boldsymbol{X}) = P(\vartheta \in \Theta_1|\boldsymbol{X}) = \int_{\Theta_1} h(\vartheta|X)\mathrm{d}\vartheta, \tag{13.3.52}$$

则贝叶斯假设检验的推断原则是

若 $\alpha_0(\boldsymbol{X}) > \alpha_1(\boldsymbol{X})$, 或 $\alpha_0(\boldsymbol{X})/\alpha_1(\boldsymbol{X}) > 1$, 接受零假设 H_0;
若 $\alpha_0(\boldsymbol{X}) < \alpha_1(\boldsymbol{X})$, 或 $\alpha_0(\boldsymbol{X})/\alpha_1(\boldsymbol{X}) < 1$, 接受备择假设 H_1;
若 $\alpha_0(\boldsymbol{X}) = \alpha_1(\boldsymbol{X})$, 问题悬而未决, 留待继续考察. (13.3.53)

比值 α_0/α_1 称为**后验概率比**：

$$\frac{\alpha_0(\boldsymbol{X})}{\alpha_1(\boldsymbol{X})} = \frac{\displaystyle\int_{\Theta_0} h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta}{\displaystyle\int_{\Theta_1} h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta} = \frac{\displaystyle\int_{\Theta_0} f(\boldsymbol{X}|\vartheta)\pi(\vartheta)\mathrm{d}\vartheta}{\displaystyle\int_{\Theta_1} f(\boldsymbol{X}|\vartheta)\pi(\vartheta)\mathrm{d}\vartheta}. \tag{13.3.54}$$

后验概率比反映了零假设 H_0 和备择假设 H_1 成立的相对可能性的大小.

需要指出, 上式中先验概率 $\pi(\vartheta)$ 在零假设 H_0 和备择假设 H_1 的参数空间 Θ_0 和 Θ_1 中可能是不同的. 设零假设 H_0 和备择假设 H_1 的先验概率分别是 π_0 和 π_1, 即

$$\pi_0 = P\left\{\vartheta \in \Theta_0\right\}, \pi_1 = P\left\{\vartheta \in \Theta_1\right\},$$

则称 π_0/π_1 为**先验概率比**.

所谓的贝叶斯因子定义为后验概率比与先验概率比的比值：

$$B(\boldsymbol{X}) = \frac{\alpha_0/\alpha_1}{\pi_0/\pi_1} = \frac{\alpha_0\pi_1}{\alpha_1\pi_0}.$$

贝叶斯因子既依赖于样本观测值 $\boldsymbol{X}$, 又依赖于先验分布 π_0 和 π_1, 两种概率比相除, 会削弱先验分布的影响, 突出样本观测值 $\boldsymbol{X}$ 的影响, 故 $B(\boldsymbol{X})$ 反映了数据 $\boldsymbol{X}$ 支持零假设 H_0 的程度.

当 $\{\Theta_0, \Theta_1\}$ 构成参数 ϑ 的完备集, 则有

$$\alpha_0(\boldsymbol{X}) + \alpha_1(\boldsymbol{X}) = \int_{\Theta_0} h\,(\theta|\boldsymbol{X})\mathrm{d}\theta + \int_{\Theta_1} h\,(\theta|\boldsymbol{X})\mathrm{d}\theta = 1.$$

故当 $\alpha_0(\boldsymbol{X}) > 0.5$, 必有 $\alpha_0(\boldsymbol{X}) > \alpha_1(\boldsymbol{X})$, 所以贝叶斯假设检验的推断原则是

$$\begin{array}{l}\text{若 } \alpha_0(\boldsymbol{X}) > 0.5, \text{ 接受零假设 } H_0;\\ \text{若 } \alpha_0(\boldsymbol{X}) < 0.5, \text{ 接受备择假设 } H_1;\\ \text{若 } \alpha_0(\boldsymbol{X}) = 0.5, \text{ 问题悬而未决, 留待继续考察.}\end{array} \tag{13.3.55}$$

现在, 我们来讨论三种检验问题.

(1) 零假设 $H_0 : \vartheta = \vartheta_0$, 备择假设 $H_1 : \vartheta = \vartheta_1$.

其中未知参数 $\vartheta \in \{\vartheta_0, \vartheta_1\}$, 且 $\vartheta_0 \neq \vartheta_1$, $\{\vartheta_0, \vartheta_1\}$ 构成 ϑ 的完备集. 这是简单原假设对简单备择假设的情形. 假定参数 ϑ 的先验分布为

$$\pi_0 = \pi(\vartheta = \vartheta_0), \quad \pi_1 = \pi(\vartheta = \vartheta_1).$$

设总体 X 的概率密度为 $f(X;\vartheta)$, 其子样测量值为 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$, 贝叶斯检验为

$$\frac{f(\boldsymbol{X}|\vartheta_1)}{f(\boldsymbol{X}|\vartheta_0)} < \frac{\pi_0}{\pi_1}, \text{接受} \quad H_0; \text{ 否则拒绝} H_0. \tag{13.3.56}$$

或

$$\frac{\pi_0 f(\boldsymbol{X}|\vartheta_0)}{\pi_1 f(\boldsymbol{X}|\vartheta_1)} > 1, \text{接受} \quad H_0; \text{ 否则拒绝} H_0. \tag{13.3.57}$$

证明：注意到当 $\vartheta = \vartheta_0$ 时 $\pi(\vartheta = \vartheta_1) = 0$, 故有

$$\alpha_0 = \int_{\Theta_0} h(\vartheta|\boldsymbol{X})\mathrm{d}\theta = h(\vartheta_0|\boldsymbol{X}) = \frac{\pi_0 f(\boldsymbol{X}|\vartheta_0)}{\pi_0 f(\boldsymbol{X}|\vartheta_0) + \pi_1 f(\boldsymbol{X}|\vartheta_1)};$$

同理, 当 $\vartheta = \vartheta_1$ 时 $\pi(\vartheta = \vartheta_0) = 0$, 故有

$$\alpha_1 = h(\vartheta_1|\boldsymbol{X}) = \frac{\pi_1 f(\boldsymbol{X}|\vartheta_1)}{\pi_0 f(\boldsymbol{X}|\vartheta_0) + \pi_1 f(\boldsymbol{X}|\vartheta_1)}.$$

代入式 (13.3.54) 得后验概率比, 与式 (13.3.53) 对比, 命题得证.

这种情形下的贝叶斯因子为

$$B(\boldsymbol{X}) = \frac{\alpha_0 \pi_1}{\alpha_1 \pi_0} = \frac{f\,(\boldsymbol{X}|\vartheta_0)}{f\,(\boldsymbol{X}|\vartheta_1)}.$$

它不依赖于先验分布, 仅依赖于样本似然比.

在经典统计的简单假设的检验问题中, Neyman — Pearson 定理告诉我们, 使检验势函数达到极大的最佳临界域由满足式 (12.2.4) 的点组成：$f(\boldsymbol{X}|\vartheta_1)/f(\boldsymbol{X}|\vartheta_0) > k$. 这与本引理中拒绝 H_0 的判断式是一致的, 并使我们对 Neyman — Pearson 定理中的临界值 k 有了新的理解, 它正是先验概率比 π_0/π_1.

(2) 零假设 $H_0:\vartheta\in\Theta_0$, 备择假设 $H_1:\vartheta\in\Theta_1$.

其中 $\vartheta\in\{\Theta_0,\Theta_1\}$, $\Theta_0\cap\Theta_1=\varnothing$, 即 $\{\Theta_0,\Theta_1\}$ 构成 ϑ 的完备集.

并假定参数 ϑ 的先验分布为

$$\pi(\vartheta)=\pi_0(\vartheta),\quad \vartheta\in\Theta_0;\quad \pi(\vartheta)=\pi_1(\vartheta),\quad \vartheta\in\Theta_1.$$

设总体 X 的概率密度为 $f(X;\vartheta)$, 其子样测量值为 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^T$, 则贝叶斯检验为

$$\frac{\displaystyle\int_{\Theta_0}\pi_0(\vartheta)f(\boldsymbol{X}|\vartheta)\mathrm{d}\vartheta}{\displaystyle\int_{\Theta_1}\pi_1(\vartheta)f(\boldsymbol{X}|\vartheta)\mathrm{d}\vartheta}>1,\text{接受}H_0;\ \text{否则拒绝}\ H_0. \tag{13.3.58}$$

该方法适用于形式如下的单侧检验或双侧检验:

零假设 $H_0:\vartheta\leqslant\vartheta_0$, 备择假设 $H_1:\vartheta>\vartheta_0$;
零假设 $H_0:\vartheta\geqslant\vartheta_0$, 备择假设 $H_1:\vartheta<\vartheta_0$;
零假设 $H_0:\vartheta_1\leqslant\vartheta\leqslant\vartheta_2$, 备择假设 $H_1:\vartheta<\vartheta_1$,$\vartheta>\vartheta_2$.

证明: 注意到当 $\vartheta\in\Theta_0$ 时 $\pi_1(\vartheta)=0$, 故有

$$\alpha_0=\int_{\Theta_0}h(\vartheta|\boldsymbol{X})\mathrm{d}\theta=P(\Theta_0|\boldsymbol{X})=\frac{\displaystyle\int_{\Theta_0}\pi_0(\vartheta)f(\boldsymbol{X}|\vartheta)\mathrm{d}\vartheta}{\displaystyle\sum_{i=0}^{1}\int_{\Theta_i}\pi_i(\vartheta)f(\boldsymbol{X}|\vartheta)\mathrm{d}\vartheta};$$

同理, 当 $\vartheta\in\Theta_1$ 时 $\pi_0(\vartheta)=0$, 故有

$$\alpha_1=P(\Theta_1|\boldsymbol{X})=\frac{\displaystyle\int_{\Theta_1}\pi_1(\vartheta)f(\boldsymbol{X}|\vartheta)\mathrm{d}\vartheta}{\displaystyle\sum_{i=0}^{1}\int_{\Theta_i}\pi_i(\vartheta)f(\boldsymbol{X}|\vartheta)\mathrm{d}\vartheta}.$$

代入式 (13.3.54) 得后验概率比, 与式 (13.3.53) 对比, 命题得证.

(3) 零假设 $H_0:\vartheta=\vartheta_0$, 备择假设 $H_1:\vartheta\neq\vartheta_0$.

其中 $\Theta_0=\{\vartheta_0\}$ 为单点集, $\Theta_1=\Theta\overline{\{\vartheta_0\}}$. 并假定参数 ϑ 的先验分布为

$$\pi(\vartheta)=\pi_0I_{\vartheta_0}(\vartheta)+\pi_1\left(\vartheta\right),$$

其中, $I_{\vartheta_0}(\vartheta)$ 是 $\vartheta=\vartheta_0$ 的示性函数, 即

$$I_{\vartheta_0}(\vartheta)=\begin{cases}1, & \vartheta=\vartheta_0;\\ 0, & \vartheta\neq\vartheta_0.\end{cases}$$

即先验分布由离散 (对 Θ_0) 与连续 (对 Θ_1) 两部分组成.

设总体 X 的概率密度为 $f(X;\vartheta)$, 其子样测量值为 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^T$, 则贝叶斯检验为

$$\frac{\pi_0 f(\boldsymbol{X}|\vartheta_0)}{\displaystyle\int_{\vartheta\neq\vartheta_0}\pi_1(\vartheta)f(\boldsymbol{X}|\vartheta)\mathrm{d}\vartheta}>1,\ \text{接受}\ H_0;\ \text{否则拒绝}\quad H_0. \tag{13.3.59}$$

证明 $\boldsymbol{X}$ 的边沿分布为

$$q(X)=\int_{\Theta}f(\boldsymbol{X}|\vartheta)\pi(\vartheta)\mathrm{d}\vartheta=\pi_0 f(\boldsymbol{X}|\vartheta_0)+q_1(\boldsymbol{X}).$$

其中,

$$q_1(\boldsymbol{X})=\int_{\vartheta\neq\vartheta_0}f(\boldsymbol{X}|\vartheta)\pi_1(\vartheta)\mathrm{d}\vartheta.$$

因此有

$$\alpha_0=\alpha_0(\boldsymbol{X})=P(\vartheta\in\Theta_0|\boldsymbol{X})=\int_{\Theta_0}h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta=\frac{\pi_0 f(\boldsymbol{X}|\vartheta_0)}{q(\boldsymbol{X})},$$

$$\alpha_1=\alpha_1(\boldsymbol{X})=P(\vartheta\in\Theta_1|\boldsymbol{X})=\int_{\Theta_1}h(\vartheta|\boldsymbol{X})\mathrm{d}\vartheta=\frac{q_1(\boldsymbol{X})}{q(\boldsymbol{X})},$$

代入式 (13.3.54) 得后验概率比, 与式 (13.3.53) 对比, 命题得证.

例 13.17 正态总体均值的贝叶斯假设检验

设 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^T$ 为正态总体 $N(\vartheta,\sigma^2)$ 的样本, ϑ,σ 均未知. 待检验的问题是

$$\text{零假设}\quad H_0: a\leqslant\vartheta\leqslant b,\quad \text{备择假设}\quad H_1:\vartheta<a\ \text{或}\ \vartheta>b.$$

其中 $a<b$, a, b 为给定常数. 用贝叶斯方法做检验.

我们采用不变先验分布 $\pi(\vartheta,\sigma)\propto 1/\sigma$, (先验分布表 13.2), 这时后验分布为

$$\begin{aligned}h(\vartheta,\sigma|\boldsymbol{X})&=\frac{L(\boldsymbol{X}|\vartheta,\sigma)\pi(\vartheta,\sigma)}{\displaystyle\int_{\Theta}L(\boldsymbol{X}|\vartheta,\sigma)\pi(\vartheta,\sigma)\mathrm{d}\vartheta\mathrm{d}\sigma}\\&=C_n\sigma^{-(n+1)}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\vartheta)^2\right],(-\infty<\vartheta<\infty,\sigma>0)\end{aligned} \tag{13.3.60}$$

这里

$$C_n=\left\{\int_0^{\infty}\int_0^{\infty}\sigma^{-(n+1)}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\vartheta)^2\right]\mathrm{d}\vartheta\mathrm{d}\sigma\right\}^{-1},$$

C_n 是与 ϑ,σ 无关、但与样本 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^T$ 有关的常数. 后面的常数 D_n,E_n,F_n 同样如此, 在对 ϑ 作统计推断时, 没有必要计算它们.

将式 (13.3.60) 对 σ 作积分, 得到 ϑ 的边沿后验分布, 为

$$\begin{aligned}&C_n\int_0^{\infty}\sigma^{-(n+1)}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\vartheta)^2\right]\mathrm{d}\sigma\\&=D_n\left[\sum_{i=1}^{n}(X_i-\vartheta)^2\right]^{-n/2}=D_n\left[S_0^2+n\left(\bar{X}-\vartheta\right)^2\right]^{-n/2}\\&=D_nS_0^{-n}\left[1+n\big(\bar{X}-\vartheta\big)^2\Big/S_0^2\right]^{-n/2}\\&=E_n\left[1+n\big(\bar{X}-\vartheta\big)^2\Big/(n-1)\,S^2\right]^{-n/2}.\end{aligned}\tag{13.3.61}$$

其中,

$$S_0^2=\sum_{i=1}^{n}\left(X-\bar{X}\right)^2=(n-1)\,S^2,$$

S^2 为样本方差, S,S_0 都与参数 ϑ,σ 无关. 令

$$\vartheta^*=\sqrt{n}\big(\vartheta-\bar{X}\big)/S,$$

ϑ^* 是 ϑ 的线性函数. 由 ϑ 的边沿后验密度可算出 ϑ^* 的边沿后验密度：

$$F_n\left(1+\frac{\vartheta^{*2}}{n-1}\right)^{-n/2},\quad -\infty<\vartheta^*<\infty.\tag{13.3.62}$$

它是自由度为 $n{-}1$ 的 t 分布的概率密度, 以 T_{n-1} 记为自由度为 $n{-}1$ 的 t 分布的累积分布函数, 于是根据式 (13.3.51), 零假设 H_0 的后验概率 α_0 为：

$$\begin{aligned}\alpha_0=\alpha_0(\boldsymbol{X})&=P(\vartheta\in H_0|\boldsymbol{X})=P(a\leqslant\vartheta\leqslant b|\boldsymbol{X})\\&=P\left(\frac{\sqrt{n}\left(a-\bar{X}\right)}{S}\leqslant\vartheta^*\leqslant\frac{\sqrt{n}\left(b-\bar{X}\right)}{S}|\boldsymbol{X}\right)\\&=T_{n-1}\left(\frac{\sqrt{n}\left(b-\bar{X}\right)}{S}\right)-T_{n-1}\left(\frac{\sqrt{n}\left(a-\bar{X}\right)}{S}\right).\end{aligned}$$

由 α_0 是否大于 0.5 来确定是否接受原假设.

例如, 设原假设为 $-1\leqslant\vartheta\leqslant 1$, $n=16$, 由样本测量值求得 $\bar{X}=0.8,S=2$, 则

$$\begin{aligned}\alpha_0=\alpha_0(\boldsymbol{X})&=T_{n-1}\left(\frac{\sqrt{n}\left(b-\bar{X}\right)}{S}\right)-T_{n-1}\left(\frac{\sqrt{n}\left(a-\bar{X}\right)}{S}\right)\\&=T_{15}\,(0.4)-T_{15}\,(-3.6)=0.655\,0-0.001\,4=0.653\,6>0.5.\end{aligned}$$

故应接受原假设.

由以上论述可见, 贝叶斯假设检验中没有“检验水平”的概念, 其方法的精神与经典的 Neyman 理论完全不同. 贝叶斯方法的好处在于：一经选定了先验分布, 剩下的只是计算问题, 不存在经典方法中寻找检验统计量及其分布函数的困难.

第 14 章　蒙特卡罗法

14.1　蒙特卡罗法的基本思想

蒙特卡罗法, 又称**统计试验法**, 是一种采用统计抽样理论近似地求解物理或数学问题的方法. 既可以求解概率问题, 也可以求解非概率问题. 蒙特卡罗法在粒子物理与核物理、宇宙线研究、原子能技术、宇航与导弹技术、运筹规划、高维数学问题、多元统计分析等许多领域中得到广泛的应用, 已经发展成为内容相当丰富的数学分支. 本章仅限于介绍该方法的基本思想和原理. 有兴趣作深入了解的读者可阅读专门的书籍和文献.

在通常的数理统计方法中, 问题的求解是通过真实的随机试验来完成的. 例如, 统计假设的检验, 利用随机过程的一组实际观测值 (随机子样的实现) 来推断随机过程总体分布的特征. 但是当利用蒙特卡罗法求解问题时, 是利用数学方法来正确地描述和模拟待求解问题的随机过程, 是一种数字模拟随机过程的方法. 对于本来不是随机性质的问题, 若用蒙特卡罗法求解, 则需人为地构造一个概率过程, 而这一过程的某个特征量恰好是问题的解. 不是依靠对过程的实际测量, 而是以对过程的模拟为基础, 这是蒙特卡罗法的基本特征.

因此, 利用蒙特卡罗法解决物理或数学问题的基本思想可归结如下：建立与待解问题相似的或相关联的概率模型或概率过程, 利用这种相似性把概率模型的某些特征量 (如随机事件的概率, 随机变量的数学期望) 与求解问题的解 (如定积分值, 微分方程的解) 联系起来, 然后对概率模型进行随机模拟或统计抽样, 这些特征量的估计值就是问题的近似解, 估计值的标准差相应于解的误差.

为了说明这些颇为抽象的叙述, 我们举一个用蒙特卡罗法计算定积分的例子. 考虑积分

$$I = \int_0^1 g(x)\mathrm{d}x, \qquad 0 \leqslant g(x) \leqslant 1. \tag{14.1.1}$$

积分 I 等于 x 在 0 与 1 之间, 曲线 $y = g(x)$ 下的面积 G(图 14.1). 为了求得面积 G, 设想在正方形 $0\leqslant x\leqslant 1$, $0\leqslant y\leqslant 1$ 内随机地投掷一个点, 该点的两个坐标在 [0, 1] 区间内均匀分布, 并且互相独立, 这样, 该点落在正方形内任一位置有相等的可能性. 于是该点落在区域 G 内的概率 p 等于 G 的面积, 即等于积分值 I. 如果用某种方法产生两个均匀分布而又相互独立的随机变量 ξ 和 η 的 N 组取样值 (ξ_i, η_i), $i = 1, 2, \cdots, N$. 对每一组 (ξ_i, η_i), 若有

$$\eta_i < f(\xi_i), \tag{14.1.2}$$

则点 (ξ_i, η_i) 落在区域 G 内, 否则落在 G 以外. 设满足不等式 (14.1.2) 的点数为 n, 则由大数定律知道, 当 $N \to \infty$ 的极限情形下, 落在区域 G 内的点数 n 与投掷的总点数 N 之比等于概率 p

$$I = p = \lim_{N\to\infty} \frac{n}{N}. \tag{14.1.3}$$

因此, 当 N 充分大, n/N 等于积分的近似值.

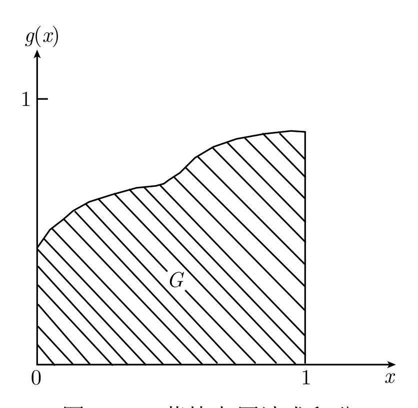


图 14.1 蒙特卡罗法求积分

由这一例子可以看到, 蒙特卡罗法解题有三个基本步骤：①构造或描述概率过程; ②实现对已知概率分布的随机抽样; ③建立与问题解对应的估计量.

本例的定积分问题是一个确定性问题, 需要人为地构造一个概率过程, 就是在 $0\leqslant x \leqslant 1$, $0\leqslant y \leqslant 1$ 的正方形内均匀、随机地投点; 描述这一随机过程的是两个独立的 [0, 1] 区间内的均匀分布, 它的抽样就是两个服从该分布的随机变量的容量 N 的子样值, 与问题解 (积分值) 对应的估计量即是频率 n/N.

由式 (14.1.3) 可知, 用蒙特卡罗法求出积分近似值的精度随着投点数 (抽样数)N 的增大而提高, 这一现象对蒙特卡罗法是具有普遍性的. 一般地说, 蒙特卡罗法解题的误差随着抽样数的增加而减小.

蒙特卡罗法的理论基础是概率论的一般定理 —— 大数定律, 因此, 它的应用范围从原则上说几乎不受什么限制. 但它与任何一种统计试验一样, 如欲获得充分可靠的结果, 需要进行大量的随机试验; 同时要使模拟试验接近真实情况, 模拟过程本身可能相当复杂, 需要很多有关的数据并进行很多次运算. 这样, 用蒙特卡罗法解题的总运算量和数据量可能很大. 只有在电子计算机出现和发展之后, 才有现实可能利用蒙特卡罗法实现大量的模拟计算. 因此, 蒙特卡罗法解题总是通过计算机实现的.

14.2　随机数的产生及检验

在上一节积分计算中, 用到 [0, 1] 区间中均匀分布的两个独立随机变量 ξ 和 η 容量 N 的子样值, 在蒙特卡罗法计算中, 一般称为服从 [0, 1] 区间均匀分布的**随机数列**, 其中的每一个体称为**随机数**. 为了求得可靠的结果, 子样容量 N 必须很大, 这就要求随机数的数量很大, 所以用蒙特卡罗法解题时, 一大部分计算工作都是随机数的运算. 此外, 随机变量的子样是相互独立并随机选取的, 随机数列的各随机数应当同样满足独立性和随机性这两个要求.

对于各种待求解的问题, 随机模型或随机过程的概率分布显然是各式各样的. 因此要求能产生服从任意概率分布的随机数列. 在后面会看到, 服从任意分布的随机数列可以用 [0, 1] 区间均匀分布的随机数列作适当的变换或舍选以后求得. 显然, 一个简便、经济、可靠而又品质良好 (即满足独立性、随机性要求) 的 [0, 1] 区间均匀分布随机数列产生器, 是蒙特卡罗法解题的关键之一.

服从一定总体分布的随机数列实际上是该总体的随机子样的实现, 而随机子样是与总体同分布的随机变量. 在后面的叙述中, 服从一定分布的随机数 (列)ξ_i 有时也看成服从同一分布的随机变量; 类似地, 随机数 (列)r_i 也可视为 [0, 1] 区间均匀分布的随机变量. 究竟表示随机变量还是随机数 (列) 由问题的叙述可以看清.

14.2.1　随机数的产生

从现在起, 我们将 [0, 1] 区间均匀分布的随机数 (列) 简称为**随机数 (列)**, 并用 r_i 标志; 服从其他分布的随机数 (列) 将标明它所服从的分布.

$r_i(i=1,2,\cdots)$ 的产生可以利用事先制订好的随机数表, 也可以用物理方法 (放射性物质的放射性, 电子线路的噪声) 产生, 但它们不适于在电子计算机中实际使用. 实际上广泛采用的是利用数学递推公式的方法, 在计算机中产生数列作为随机数列.

随机数列产生的递推公式有如下形式：

$$r_{n+k}=T(r_n,r_{n+1},\cdots,r_{n+k-1}), \tag{14.2.1}$$

T 是某个函数, 给定初值 r_1, r_2, $\cdots$, r_k, 按上式可确定 $r_{n+k}(n=1,2,\cdots)$, 构成随机数列. 经常应用的是 $k=1$ 的情况, 递推公式简化为

$$r_{n+1}=T(r_n). \tag{14.2.2}$$

对于给定的初值 r_1, 由上式可确定随机数列 $r_{n+1}(n=1,2,\cdots)$.

递推公式产生的随机数列存在两方面的问题：

(1) 递推公式和初始值 r_1, $\cdots$, r_k 确定后, 整个随机数列就唯一地确定了. 故不满足随机数之间相互独立的要求.

(2) 随机数列是按确定性的算法计算出来的, 而且在电子计算机上所能表示的 [0, 1] 区间内的数只是有限多个 (由计算机字长所限定). 因此, 递推到了一定的次数, 同一个数字总会出现第二次, 而此后就出现周期性的重复现象. 因此, 这样的数列不符合随机性 (对均匀分布也就是均匀性) 的要求.

由于这两个原因, 用数学递推公式产生的随机数通常称为**伪随机数**. 伪随机数的这两个缺点不可能从根本上加以改变. 但只要递推公式选得比较好, 随机数的相互独立性可以近似地得到满足, 重复的周期又可以足够地长, 使得在用蒙特卡罗法解题时, 实际使用的随机数列长度小于出现重复的周期长度, 这样, 第二个问题实际上不存在, 保证了随机数列在 [0, 1] 区间内的均匀性. 这种方法的最大优点是占用计算机内存小、产生速度快、可以复算、不受计算机型号的限制, 因此是电子计算机产生随机数列的最主要方法.

递推公式产生伪随机数列的方法有好多种, 这里只介绍使用较广泛的同余法中的三种, 它们能产生周期长、统计性质好的伪随机数列.

(a) 乘同余法

递推公式为

$$x_{n+1} \equiv a x_n (\mathrm{mod} M), \qquad r_{n+1} = \frac{x_{n+1}}{M}, \tag{14.2.3}$$

其中 a 是乘因子, M 为正整数 (称为模数), 第一个式子称为模数为 M 的同余式, mod 是 modula 的缩写, 该式的含义表示 ax_n 除以 M 后所得的余数等于 x_{n+1}. 给定任意初值 x_0, 可算得数列 x_1, x_2, $\cdots$, 除以 M 后得到所需要的伪随机数列 r_1, r_2, $\cdots$.

从数列的构造过程可知, 每一个 x_n(因而 r_n) 最多有 M 个相异值, 即

$$0 \leqslant x_n \leqslant M, \qquad 0 \leqslant r_n \leqslant 1.$$

所以数列 $\{r_n\}$ 的最大容量 $L \leqslant M$. 在计算机中, M 的取值通常是 2^k, k 为字长 (二进制数的最大可能有效位数).

乘同余法产生的数列的独立性和分布的均匀性取决于参数 x_0、a 的选择.

(b) 乘加同余法

递推公式为

$$x_{n+1} \equiv a x_n + C (\mathrm{mod} M), \qquad r_{n+1} = \frac{x_{n+1}}{M}, \tag{14.2.4}$$

C 是非负整数. 数列 $\{r_{n+1}\}$ 的最大容量 $L \leqslant M$. 适当选取参数 C 可改善伪随机数列的统计性质.

(c) 加同余方法

递推公式为

$$x_{n+2} = x_n + x_{n+1} (\mathrm{mod} M), \qquad r_{n+2} = \frac{x_{n+2}}{M}. \tag{14.2.5}$$

由任意初始值 x_0、x_1 可产生伪随机数列 $\{r_n\}, n=2,3,\cdots$.

加同余方法产生的伪随机数列的最大容量在一般情况下来考虑是很困难的. 对于 $M=2^k, x_0=x_1=1$ 的特殊情形, 最大容量为 $1.5M$.

14.2.2 随机数的统计检验

不管用什么方法产生伪随机数列, 它们能否作为随机数列使用, 最终要靠统计检验来确定. 检验的内容主要是随机数分布参数是否与理论分布一致, 随机数是否有较好的均匀性、独立性和连贯性, 它们之间既有一定的差别, 又有一定的联系. 例如均匀性检验方法虽然侧重于随机数均匀性的检验, 但同时也检验了分布参数、独立性等, 其他检验方法也有类似的情形. 检验同一个随机性质可以采用不同的统计检验方法, 多做几种统计检验可以保证所产生的随机数有较好的统计性质. 此外, 对随机数列的统计检验应当按照问题的性质有所侧重. 如果问题主要要求随机数的均匀性, 如一维定积分的计算, 则应侧重均匀性的检验; 反之, 若问题的解容易受 r_i 之间相关性质或出现顺序的影响, 则应着重作独立性和连贯性检验.

1) 参数检验

随机数的参数检验是检验所产生的随机数列的分布参数是否与 [0, 1] 区间均匀分布随机变量 r 的相应分布参数一致. 在产生 N 个随机数 $r_1, r_2, \cdots, r_N$ 之后, 因为它们可看成总体子样的 N 个观测值, 故可以构成各阶子样矩 (见 6.2 节)

$$\hat{m}_k=\frac{1}{N}\sum_{i=1}^{N}r_i^k,\qquad k=1,\ 2,\ \cdots. \tag{14.2.6}$$

若 H_0 为真, 各阶子样矩的期望值和方差的理论值为

$$m_k=\frac{1}{k+1},\qquad \sigma_{k,N}^2=\left(\frac{1}{2k+1}-m_k^2\right)\frac{1}{N}. \tag{14.2.7}$$

根据同分布中心极限定理, 统计量

$$Z_{k,N}=\frac{\hat{m}_k-m_k}{\sigma_{k,N}}, \tag{14.2.8}$$

渐近地服从标准正态分布 $N(0, 1)$. 按照给定的显著性水平检验假设 H_0, 就是由累积标准正态分布表查出满足

$$\Phi(Z_{\alpha/2})=1-\alpha/2 \tag{14.2.9}$$

的值 $Z_{\alpha/2}$, 当 N 足够大, 若观测到的 $Z_{k,N}$ 大于 $Z_{\alpha/2}$, 则在显著性水平 α 上拒绝 H_0; 反之, 则可以认为原假设成立, 即随机数的参数 (各阶矩) 与 [0, 1] 区间均匀分布随机变量的参数一致.

2) 均匀性检验

检验随机数的均匀性, 即随机数列 r_1, r_2, $\cdots$ 是否在 [0, 1] 区间均匀分布. 可通过频率分布、累积频率分布的测定值与理论值之间的比较来实现.

(a) 频率检验. 将 [0, 1] 区间划分为 k 个等长的子区间, 每个子区间长度为 $1/k$. 设产生了 N 个伪随机数 $r_1, r_2, \cdots, r_N$. 落在第 j 个子区间内的伪随机数个数记为 $n_j(j=1,2,\cdots,k)$, 称为经验频数. 显然

$$\sum_{j=1}^{k} n_j = N$$

成立.

如零假设 H_0 为真, 这 N 个随机数 (即 [0, 1] 区间均匀分布的容量 N 的子样观测值) 落在任一子区间内的概率 p_j 等于子区间的长度

$$p_j = 1/k.$$

故落在任一子区间内的随机数个数的理论值 (理论频数) 为

$$m_j = Np_j = N/k, \quad j=1,2,\cdots,k.$$

按照 12.4.2 节的讨论, 统计量

$$\chi^2 = \sum_{j=1}^{k} \frac{(n_j - m_j)^2}{m_j} \tag{14.2.10}$$

渐近地服从自由度 $k-1$ 的 χ^2 分布. 因此, 对于给定的显著水平 α, 可利用皮尔逊检验来确定随机数列是否满足均匀分布的零假设 H_0.

(b) 累积频率检验. 12.4.5 节介绍的柯尔莫哥洛夫检验法可以用来检验随机数分布的均匀性. 该检验方法利用统计量

$$D_N = \max_{-\infty<x<+\infty} \left| F_N(x) - F_0(x) \right|,$$

其中 $F_n(x)$ 是随机变量 X 的子样分布函数, $F_0(x)$ 是零假设的累积分布函数.

在现在的情形中, 将所产生的 N 个随机数按数值递增的次序排列

$$r_1, \ r_2, \ \cdots, \ r_N.$$

由 6.1 节的讨论可知, 子样分布函数 $F_N(r)$ 在 r_i 点的数值为

$$F_N(r_i) = i/N.$$

这也就是随机数列的累积频率. 而零假设即是 [0, 1] 区间的均匀分布, 其累积分布在 r_i 点的数值是

$$F_0(r_i) = r_i.$$

所以令 $n(x)$ 是 $r_1, r_2, \cdots, r_N$ 中满足

$$r_i < x$$

的随机数个数, 则统计量 D_N 可表示为

$$D_N = \max_{0<x<1} \left| \frac{n(x)}{N} - x \right|. \tag{14.2.11}$$

然后用 12.4.5 节介绍的检验方法确定, 在给定的显著性水平 α 下接收还是拒绝零假设 H_0.

3) 独立性检验

(a) 相关系数检验. 设产生了 N 个随机数

$$r_1, \ r_2, \ \cdots, \ r_N.$$

各随机数之间相互独立的必要条件是它们的相关系数等于 0. 若前后相距为 j 的随机数之间的相关系数记为 ρ_j, 按相关系数的定义, ρ_j 的估计量为

$$\hat{\rho}_j = \left[\frac{1}{N-j} \sum_{i=1}^{N-j} r_i r_{j+i} - (\bar{r})^2 \right] / s^2, \qquad j = 1, \ 2, \ \cdots, \tag{14.2.12}$$

其中,

$$s^2 = \frac{1}{N} \sum_{i=1}^{N} \left(r_i - \frac{1}{2} \right)^2.$$

对充分大的 N(如 $N - j > 50$), 当原假设 H_0: $\rho_j = 0$ 为真, 统计量

$$u = \hat{\rho}_j \sqrt{N-j} \tag{14.2.13}$$

渐近地服从标准正态分布. 因而 u 可作为随机数列 $r_1, r_2, \cdots, r_N$ 独立性检验的检验统计量.

(b) 联列表独立性检验. 在 12.6 节中我们讨论了二维随机变量的联列表独立性检验, 该方法也可用来对随机数的独立性进行检验.

在 xy 平面上将 $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ 的正方形分成 $J \times K$ 个矩形, 用任意一种方法将伪随机数列 $\{r_i\}(i = 1, 2, \cdots, N)$ 两两组成二维空间上的点列 $\{r_{i,1}, r_{i,2}\}(i = 1, 2, \cdots, n)$, n 为小于等于 $N/2$ 的最大整数. 这些点中落入第 j、k 个矩形中的数目记为 n_{jk}, $(j = 1, \cdots, J;\ k = 1, \cdots, K)$. 记

$$n_{j\cdot} = \sum_{k=1}^{K} n_{jk}, \quad n_{\cdot k} = \sum_{j=1}^{J} n_{jk},$$

可有以下联列表 (表 14.1). 显然

$$\sum_{j=1}^{J}\sum_{k=1}^{K} n_{jk} = \sum_{j=1}^{J} n_{j\cdot} = \sum_{k=1}^{K} n_{\cdot k} = n.$$

表 14.1 二维随机变量的联列表

j \ k	1	2	$\cdots$	K	合计 $n_{j\cdot}$
1	n_{11}	n_{12}	$\cdots$	n_{1K}	$n_{1\cdot}$
2	n_{21}	n_{22}	$\cdots$	n_{2K}	$n_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
J	n_{J1}	n_{J2}	$\cdots$	n_{JK}	$n_{J\cdot}$
合计 $n_{\cdot k}$	$n_{\cdot 1}$	$n_{\cdot 2}$	$\cdots$	$n_{\cdot K}$	N

根据 12.6 节的论证, 统计量

$$\chi^2 = n\sum_{j=1}^{J}\sum_{k=1}^{K} \frac{\left(n_{jk} - \dfrac{n_{j\cdot}n_{\cdot k}}{n}\right)^2}{n_{j\cdot}n_{\cdot k}} \tag{14.2.14}$$

渐近地服从自由度 $(J-1)(K-1)$ 的 χ^2 分布. 利用 χ^2 检验法便可确定在给定显著性水平 α 上, 零假设 H_0(相互独立的随机数列) 是否被接受.

(c) 多维频率检验. 将伪随机数列 $r_1, r_2, \cdots, r_N$ 用任意一种方法, 每 s 个随机数组成 s 维空间上的一个点, 构成一个点列 $\{r_{i,1}, r_{i,2}, \cdots, r_{i,s}\}(i=1,2,\cdots,n)$, n 是小于等于 N/s 的最大整数. 把 s 维空间上边长为 1 的超立方体分成 k 个子区域, 令 n_k 为这 n 个点中落入 k 子区域的点数 (经验频数). 如果 H_0 为真, 即 r_1, $r_2, \cdots, r_N$ 是 [0, 1] 区间均匀分布的容量 N 的独立随机子样的观测值, 那么属于 k 子区域的理论频数为

$$m_k = np_k, \qquad k = 1,\ 2,\ \cdots,\ K,$$

其中 p_k 为随机数属于 k 子区域的概率, 等于 k 子区域的 “体积”. 此时, 统计量

$$\chi^2 = \sum_{k=1}^{k} \frac{(n_k - m_k)^2}{m_k} \tag{14.2.15}$$

渐近地服从自由度 $K-1$ 的 χ^2 分布, 可以用 χ^2 检验法来检验伪随机数列的独立性假设 H_0 是否成立.

(d) 多维矩检验. 利用 (c) 描述的方法将伪随机数列 $r_1, r_2, \cdots, r_N$ 构成 s 维空间上的 n 个点 $\{r_{i,1}, r_{i,2}, \cdots, r_{i,s}\}(i=1,2,\cdots,n)$. 观测值的多维矩 (子样多维

矩) 是

$$\hat{m}_{k_1\cdots k_s}=\frac{1}{n}\sum_{i=1}^{n}r_{i,1}^{k_1}\cdot r_{i,2}^{k_2}\cdot\cdots\cdot r_{i,s}^{k_s},$$

对 [0, 1] 区间均匀分布的容量 N 的独立子样, 多维矩及其方差的理论值为

$$m_{k_1\cdots k_s}=\frac{1}{(k_1+1)(k_2+1)\cdots(k_s+1)},$$

$$\sigma_{k_1\cdots k_s}^2=\frac{1}{n}\left\{\frac{1}{(2k_1+1)(2k_2+1)\cdots(2k_s+1)}-m_{k_1k_2\cdots k_s}^2\right\}.$$

根据中心极限定理, 统计量

$$Z_{k_1\cdots k_s}=\frac{\hat{m}_{k_1\cdots k_s}-m_{k_1\cdots k_s}}{\sigma_{k_1\cdots k_s}}\tag{14.2.16}$$

渐近地服从标准正态分布 $N(0,1)$. 可用类似于矩检验的方法来决定在给定的显著性水平上接受还是拒绝零假设 H_0.

4) 连贯性检验. 随机数的连贯性检验是按照随机数出现的先后顺序, 检验它的连贯现象是否异常.

将随机数列 $r_1, r_2, \cdots, r_N$ 按某种规律分成两类, 分别称为 a 类和 b 类, 属于 a 类的概率为 p, 属于 b 类的概率为 $q=1-p$. 例如, $r_i\leqslant p$ 的 r_i 属于 a 类, $r_i>p$ 的 r_i 属于 b 类, 即是一种分法. 按随机数出现的先后顺序进行排列：

$$\underbrace{aabbbaaababb\cdots}_{N\text{个}}$$

相连的同类元素构成游程 (见 12.7.2 节). 令 m、n 分别为 a、b 类元素的个数, 显然 $N=m+n$. 总游程数记为 R, 它的分布为

$$p(R=2k)=2\begin{pmatrix}m-1\\k-1\end{pmatrix}\begin{pmatrix}n-1\\k-1\end{pmatrix}p^mq^n,$$

$$p(R=2k+1)=\left[\begin{pmatrix}m-1\\k\end{pmatrix}\begin{pmatrix}n-1\\k-1\end{pmatrix}+\begin{pmatrix}m-1\\k-1\end{pmatrix}\begin{pmatrix}n-1\\k\end{pmatrix}\right]p^mq^n.$$

R 的数学期望和方差为

$$\begin{aligned}E(R)&=p^2+q^2+2Npq,\\V(R)&=4Npq(1-3pq)-2pq(3-10pq).\end{aligned}\tag{14.2.17}$$

当 N 充分大, 统计量 Z 渐近地服从标准正态分布

$$Z=[R-E(R)]/\sqrt{V(R)}.\tag{14.2.18}$$

因此, Z 可作为检验随机数列连贯性的检验统计量.

常用 $p=q=1/2$, 这时

$$E(R)=\frac{N}{2}+\frac{1}{2},\qquad V(R)=\frac{N-1}{4}.$$

因此

$$Z=\frac{2R-N-1}{\sqrt{N-1}}\sim N(0,1). \tag{14.2.19}$$

14.3 任意随机变量的随机抽样

随机变量的随机抽样指的是由该变量的总体分布产生简单随机子样. 在用蒙特卡罗法解题时, 经常会遇到具有不同分布的各种随机变量, 要求产生对应于该随机变量的随机子样, 也即随机数列. 这一步骤称为对该随机变量的随机模拟或随机抽样.

上一节我们讨论了随机数的独立性和均匀性 (即同分布) 检验. 下面我们将会见到, 任意分布的随机数列可由 [0, 1] 区间均匀分布的随机数列经过变换或舍选得到. 只要 r_i 满足均匀且相互独立的要求, 则由它所产生的任何分布的随机数列 (简单随机子样) 相互独立且与总体同分布.

14.3.1 直接抽样方法

(1) 离散随机变量

设 X 为离散随机变量, 它取值 $x_1, x_2, \cdots, x_n$ 的概率用 $p_1, p_2, \cdots, p_n$ 表示. 其累积概率可表示为

$$F(x)=\sum_{x_i<x}p_i,$$

归一性要求

$$F(x\geqslant x_n)=\sum_{i=1}^{n}p_i=1.$$

显然, $F(x)$ 的取值在 [0, 1] 区间内.

由图 14.2 可见, 当 $F(x)$ 的值在

$$\sum_{i=1}^{i^*-1}p_i,\quad \sum_{i=1}^{i^*}p_i$$

之间, 随机变量取值为 x_{i^*}. 因为任意随机变量的累积分布 $F(x)$ 的 pdf 为 $[0,1]$ 区间的均匀分布, 因此, 抽取一个 [0, 1] 区间均匀分布随机数 r, 当

$$\sum_{i=1}^{i^*-1}p_i<r\leqslant\sum_{i=1}^{i^*}p_i \tag{14.3.1}$$

成立, 则随机变量应取值 x_{i^*}, 即

$$\xi = x_{i^*}$$

是随机变量 X 的抽样值. 这就是离散随机变量的直接抽样方法.

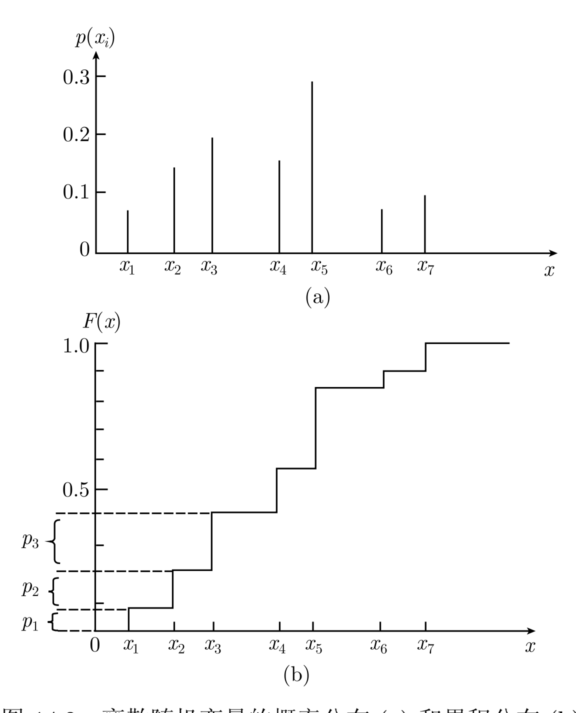


图 14.2　离散随机变量的概率分布 (a) 和累积分布 (b)

在具体对某种离散随机变量进行抽样时, p_i 不一定有简单的表达式, 而利用分布本身的某些特点往往可得到简单的抽样方法.

下面讨论离散随机变量随机抽样的几个具体例子：

(a) 粒子衰变末态的随机抽样. 设粒子 a 有三种衰变方式, 每一种衰变方式的概率 (称为分支比) 如下：

$$\begin{aligned} a \rightarrow b_1 + c_1 \quad & p_1 = 0.5, \\ b_2 + c_2 \quad & p_2 = 0.3, \\ b_3 + c_3 \quad & p_3 = 0.2, \end{aligned}$$

随机抽样的方式为：产生一个随机数 r, 若

$$\begin{aligned} &0 < r \leqslant 0.5, && \text{末态抽样为} b_1 + c_1, \\ &0.5 < r \leqslant 0.5 + 0.3, && \text{末态抽样为} b_2 + c_2, \\ &0.5 + 0.3 < r \leqslant 1, && \text{末态抽样为} b_3 + c_3. \end{aligned}$$

对于存在更多衰变末态的场合可依此类推.

(b) 伯努利试验的抽样. 伯努利试验的结果只有两种：成功和失败. 定义一个随机变量 X, 当试验成功, 取值 $x=1$; 试验失败, 取值 $x=0$. 这个变量服从的分布为伯努利分布. 设成功的概率为 p, 伯努利分布的随机抽样方法为：产生一个随机数 r,

$$\begin{aligned} &r<p, \quad \xi=1, \\ &r \geqslant p, \quad \xi=0. \end{aligned} \tag{14.3.2}$$

则 ξ 为服从伯努利分布的抽样值.

(c) 二项分布的抽样. 二项分布是 n 次独立的伯努利试验中, 事件 A(试验成功) 的次数等于 k 的概率分布, 它可表示为

$$p(n,k)=C_n^k p^k(1-p)^{n-k}, \quad k=0,1,\cdots,n.$$

当所有的 $p(n,\, k)(k=0,1,\cdots,n)$ 为已知或容易计算时, 可直接根据式 (14.3.1) 描述的方法抽样. 但当 n 很大时, 求和计算和 $p(n,\, k)$ 的计算都很繁杂. 以下方法实现抽样更为简单易行; 产生 n 个随机数 $r_1, r_2, \cdots, r_n$, 统计满足

$$r_i<p, \qquad i=1,2,\cdots,n \tag{14.3.3}$$

的随机数个数 ξ, 由式 (14.3.2) 可知, ξ 是 n 次伯努利试验中成功的次数, 因此, ξ 就是二项分布的抽样值.

(d) 泊松分布的抽样. 泊松分布的概率表达式为

$$P(k;\ \lambda)=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}, \qquad k=0,1,2,\cdots.$$

显然, 用式 (14.3.1) 进行抽样的计算相当繁杂.

根据泊松定理, 若进行 n 次独立的伯努利试验, 每次试验成功的概率为 p, 在 $n\to\infty$, $p\to 0$, $np=\lambda$ 条件下, 试验成功为 k 次的概率趋近于均值为 λ 的泊松分布. 可以根据这一原理来构造泊松分布的抽样过程. 选取足够大的 n, 使 $p=\lambda/n$ 相当小, 如 $p=\lambda/n\leqslant 0.1$, 产生 n 个随机数 $r_1, r_2, \cdots, r_n$, 其中满足

$$r_i<p, \qquad i=1,2,\cdots,n$$

的个数 ξ 表示伯努利试验成功的次数, 故 ξ 近似地为泊松分布的抽样值. n 越大, 近似程度越好.

(2) 连续随机变量 (反函数法)

设任意连续随机变量 Y 的概率密度为 $g(y)$, 累积分布函数为 $G(y)$, 在 4.7 节中已经证明

$$x=G(y)=\int_{-\infty}^{y} g(t)\mathrm{d}t$$

为 [0, 1] 区间中均匀分布的随机变量. 若 $G(y)$ 的反函数存在, 上式可写成

$$y = G^{-1}(x).$$

所以若令 ξ_i 为随机变量 Y 的随机子样 (随机数列), r_i 为 [0, 1] 区间均匀分布的随机子样 (随机数列), 则应有

$$r_i = G(\xi_i) = \int_{-\infty}^{\xi_i} g(y)\mathrm{d}y, \tag{14.3.4}$$

写成反函数的形式为

$$\xi_i = G^{-1}(r_i). \tag{14.3.5}$$

利用这两个公式可从 r_i 求出任意连续随机变量的随机数列 ξ_i. 例如, 求区间 $[a, b]$ 上均匀分布的随机数列 ξ_i, 根据式 (14.3.4)

$$r_i = \int_a^{\xi_i} \frac{1}{b-a}\mathrm{d}y = \frac{\xi_i - a}{b-a},$$

故有 $\xi_i = r_i(b-a) + a$.

又如指数分布 $f(x) = \lambda \mathrm{e}^{-\lambda x}(x \geqslant 0)$ 的随机数 ξ_i, 根据式 (14.3.4),

$$r_i = \int_0^{\xi_i} \lambda \mathrm{e}^{-\lambda x}\mathrm{d}x = 1 - \mathrm{e}^{-\lambda \xi_i},$$

求得

$$\xi_i = -\frac{1}{\lambda}\ln(1 - r_i).$$

注意到 r_i 与 $1-r_i$ 都是 [0, 1] 区间均匀分布随机数, 故上式可简化为

$$\xi_i = -\frac{1}{\lambda}\ln r_i.$$

14.3.2 直接抽样方法的推广 —— 变换抽样

直接抽样方法的实质在于由 [0, 1] 区间均匀分布随机数 r_i 经过某种变换得到任意分布的随机数. 对于连续随机变量, 当该变量的分布函数的反函数存在, 这个变换就是反函数.

这种方法可以推广到更一般的情况, 如已知某个随机变量 X 的概率密度 $f(x)$ 及其随机数 η_i, 则任意随机变量的随机数 ξ_i 可通过对 η_i 的某种变换求得.

设随机变量

$$Y = Y(X) \tag{14.3.6}$$

是 X 的一一对应的变换, 2.3 节已经证明, Y 的概率密度为

$$g(y) = \left|\frac{\mathrm{d}x(y)}{\mathrm{d}y}\right| f(x(y)).$$

随机变量 Y(或概率分布 $g(y)$) 的随机数 ξ_i 可由下式求得：

$$\xi_i = y(\eta_i) \tag{14.3.7}$$

式 (14.3.4) 和式 (14.3.5) 只是该式的特例. 因为若令 X 为 [0, 1] 区间均匀分布的随机变量, $f(x) = 1$; 任意随机变量 Y 的概率密度令为 $g(y)$, 则有

$$g(y) = \left|\frac{\mathrm{d}x(y)}{\mathrm{d}y}\right|, \qquad \mathrm{d}x = g(y)\mathrm{d}y = \mathrm{d}G(y),$$

两边求积分, 得

$$x = G(y) = \int_{-\infty}^{y} g(y)\mathrm{d}y,$$

于是有

$$y = G^{-1}(x),$$

对照式 (14.3.6) 和式 (14.3.7) 即得 $\xi_i = G^{-1}(r_i)$, 这正是式 (14.3.5).

由式 (14.3.7), 可从任一已知分布的随机数 η_i 求出任意连续分布的随机数 ξ_i. 例如由 6.3.2 节知, 若 X_1, X_2, $\cdots$, X_n 为相互独立的标准正态随机变量, 则

$$\chi^2 = \sum_{i=1}^{n} X_i^2$$

服从 $\chi^2(n)$ 分布. 令 δ_i 为服从 $N(0, 1)$ 的随机数, 则 $\chi^2(n)$ 的随机数 ξ 可表示为

$$\xi = \sum_{i=1}^{n} \delta_i^2.$$

由 $\chi^2(n)$ 的可加性, 即

$$\chi^2(n_1 + n_2) = \chi^2(n_1) + \chi^2(n_2)$$

立即可得 $\chi^2(n_1 + n_2)$ 的随机数 ξ_{1+2} 为 $\chi^2(n_1)$、$\chi^2(n_2)$ 随机数 ξ_1、ξ_2 之和

$$\xi_{1+2} = \xi_1 + \xi_2.$$

其他有可加性的分布的随机数有类似的公式.

这种抽样方法对于多维随机变量同样适用. 3.7 节中已经证明, 若随机变量 $\boldsymbol{X} = \{X_1, X_2, \cdots, X_n\}$ 的概率密度函数为

$$f(\boldsymbol{x}) = f(x_1, x_2, \cdots x_n),$$

随机变量 $\boldsymbol{Y} = \{Y_1, Y_2, \cdots, Y_n\}$ 是 $\boldsymbol{X}$ 的函数

$$Y_i = Y_i(\boldsymbol{X}), \qquad i = 1, 2, \cdots, n, \tag{14.3.8}$$

则随机变量 $\boldsymbol{Y}$ 的概率密度可表示为

$$g(\boldsymbol{y}) = f(\boldsymbol{x}) \left| J\left(\frac{\boldsymbol{x}}{\boldsymbol{y}}\right) \right|.$$

若 $\boldsymbol{X}$ 的随机数向量用 $\boldsymbol{\eta} = \{\eta_1, \eta_2, \cdots, \eta_n\}$ 表示, 则 $\boldsymbol{Y}$ 的随机数向量 $\boldsymbol{\xi} = \{\xi_1, \xi_2, \cdots, \xi_n\}$ 由下式求出:

$$\xi_i = y_i(\boldsymbol{\eta}), \qquad i = 1, 2, \cdots, n. \tag{14.3.9}$$

例如, 两个相互独立的正态 $N(0,\ 1)$ 的随机数 ξ_1、ξ_2 可按下式求得:

$$\begin{aligned} \xi_1 &= (-2\ln r_1)^{1/2}\cos(2\pi r_2), \\ \xi_2 &= (-2\ln r_1)^{1/2}\sin(2\pi r_2). \end{aligned} \tag{14.3.10}$$

证明如下:

将上式中 r_1、r_2、ξ_1、ξ_2 看作随机变量, 可解得

$$r_1 = \mathrm{e}^{-\frac{\xi_1^2+\xi_2^2}{2}}, \quad r_2 = \frac{1}{2\pi}\arctan\frac{\xi_2}{\xi_1}.$$

变换的雅可比行列式是

$$J\left(\frac{\boldsymbol{r}}{\boldsymbol{\xi}}\right) = J\left(\frac{r_1, r_2}{\xi_1, \xi_2}\right) = \begin{vmatrix} \dfrac{\partial r_1}{\partial \xi_1} \dfrac{\partial r_1}{\partial \xi_2} \\ \dfrac{\partial r_2}{\partial \xi_1} \dfrac{\partial r_2}{\partial \xi_2} \end{vmatrix}$$

$$= \begin{vmatrix} -\xi_1 \mathrm{e}^{-\frac{\xi_1^2+\xi_2^2}{2}} & -\xi_2 \mathrm{e}^{-\left(\frac{\xi_1^2+\xi_2^2}{2}\right)} \\ \dfrac{1}{2\pi}\dfrac{-\dfrac{\xi_2}{\xi_1^2}}{1+\left(\dfrac{\xi_2}{\xi_1}\right)^2} & \dfrac{1}{2\pi}\dfrac{\dfrac{1}{\xi_1}}{1+\left(\dfrac{\xi_2}{\xi_1}\right)^2} \end{vmatrix}$$

$$= -\frac{1}{2\pi}\mathrm{e}^{-\frac{\xi_1^2+\xi_2^2}{2}} = -\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{\xi_1^2}{2}}\right)\cdot\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{\xi_2^2}{2}}\right).$$

又 $\{r_1, r_2\}$ 的联合概率密度为

$$f(r_1, r_2) = 1,$$

因此, $\boldsymbol{\xi}$ 的联合概率密度为

$$g(\boldsymbol{\xi}) = f(\boldsymbol{r}) \cdot \left| J\left(\frac{\boldsymbol{r}}{\boldsymbol{\xi}}\right) \right| = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{\xi_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{\xi_2^2}{2}},$$

即等于两个 $N(0,\ 1)$ 概率密度的乘积. 因此, ξ_1, ξ_2 是相互独立的正态 $N(0,\ 1)$ 随机数.

14.3.3 舍选抽样方法

对于连续随机变量, 直接抽样方法在一些情况下会遇到困难. 第一, 许多分布无法用解析函数给出; 第二, 有些分布函数的反函数不存在或难以求出; 第三, 反函数虽然可以求出, 但运算量太大. 在这些情形下, 比较适用的是舍选抽样方法.

舍选法的实质是从随机数列 $r_i(i=1,2,\cdots)$ 中按一定的舍选规则选出其中的一部分, 使之成为具有给定分布的随机数列.

设随机变量 X 的取值域为 $[a, b]$, 概率密度 $f(x)$ 为有界函数, 其极大值用 M 表示

$$f(x) \leqslant M,$$

则该分布的舍选抽样可用图 14.3 表示.

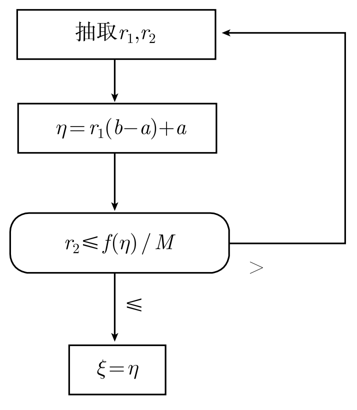


图 14.3 舍选抽样方法

$a \leqslant x \leqslant b,\ 0 \leqslant y \leqslant M$

框图含义如下：η 是 $[a, b]$ 区间内均匀分布随机数, 考察 $r_2 \leqslant f(\eta)/M$ 是否成立. 若成立, 则 η 取为随机变量 X 的随机数 ξ; 否则, η 被舍弃. 重复以上过程可求得 X 的随机数列 $\{\xi_i\}$.

该方法的实际内容是在区域内产生均匀、相互独立的随机点列 $(\eta_1, Mr_2)(\eta_3, Mr_4), \cdots, (\eta_{2N-1}, Mr_{2N})$, 抛弃在 $f(x)$ 曲线之上的所有点, 保留曲线 $f(x)$ 下的所有点, 从而形成

$$a \leqslant x \leqslant b, \qquad 0 \leqslant y \leqslant f(x)$$

区域内均匀的相互独立的随机点列 $(\xi_1, y_1), (\xi_2, y_2), \cdots$, 其中的 $\xi_1, \xi_2, \cdots$ 便是总体分布 $f(x)$ 的随机子样.

上面假定 $[a, b]$ 是有限区间. 对于无限区间的情形, 可进行截尾处理, 即选择有限区间 $[a, b]$, 满足

$$\int_a^b f(x)\mathrm{d}x \geqslant 1-\varepsilon. \tag{14.3.11}$$

只要 ε 足够小, 就可以应用上述方法而使计算误差满足要求. 这种对无限区间的截尾处理在本章讨论的各种抽样方法都适用.

我们来讨论舍选抽样的抽样效率. 如果选出某特定分布的一个随机数 ξ 平均地需要 n 个随机数 r_i, 则抽样效率

$$E = \frac{1}{n}.$$

一个好的抽样方法的抽样效率应当尽可能接近于 1. 在舍选法中, 假定产生了 n 对 r_1, r_2(共 $2n$ 个 r_i), 其中满足不等式 $r_2 \leqslant f(\eta)/M$ 的 η 的个数 m 与全部 η 的个数 n 之比, 等于曲线 $f(x)$ 下的面积与总面积 $(b-a)M$ 之比 (图 14.4)

$$\frac{m}{n} = \frac{\displaystyle\int_a^b f(x)\mathrm{d}x}{(b-a)M} = \frac{1}{(b-a)M}$$

因此, 舍选法的抽样效率为

$$E = \frac{m}{2n} = \frac{1}{2M(b-a)}. \tag{14.3.12}$$

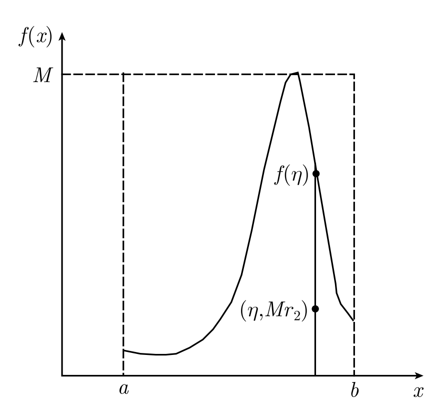


图 14.4　舍选法的抽样效率

显然, $f(x)$ 的变化越是平缓, M 值就越小, 抽样效率就越高; 反之, 若 $f(x)$ 变化很大, 而且函数形状窄而尖锐, 则 M 值很大, 抽样效率很低. 所以该方法适用于 $f(x)$ 变化不大的随机变量.

求得一个特定分布的随机数 ξ_i 所需的计算机平均机时称为抽样费用. 显然抽样费用越小越有利, 它是衡量抽样过程合理性的综合指标. 它不仅取决于抽样效率, 还取决于产生随机数 r 的运算时间以及抽样过程中的其他运算. 当产生随机数 r 的方法固定之后, 产生 r 的运算时间为一常数, 抽样费用由其他两个因素决定.

在舍选法中, 如果函数 $f(x)$ 的形式很复杂, 计算费用也就相应地增大.

14.3.4 利用极限定理抽样

第 5 章同分布的中心极限定理告诉我们, 当 $n \to \infty$ 时, n 个独立、同分布随机变量 $X_1, X_2, \ldots, X_n$ 构成的随机变量 Y 有如下分布：

$$Y = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \sim N(0,1),$$

其中 μ、σ^2 是 X 的数学期望和方差. 利用这一性质可从随机数 r_i 构成标准正态分布的随机数 ξ. 由于 [0, 1] 区间均匀分布的随机变量的 μ 和 σ^2 分别为

$$\mu = \frac{1}{2}, \qquad \sigma^2 = \frac{1}{12},$$

因此 $N(0, 1)$ 的随机数可表示为

$$\xi = \left(\sum_{i=1}^{n} r_i - \frac{n}{2}\right) \Big/ \sqrt{\frac{n}{12}}. \tag{14.3.13}$$

为了保证精确度, n 应当足够大, 一般情形下 $n \geqslant 10$ 时上式的正态近似已经相当好. 取 $n = 12$, 式 (14.3.13) 形式特别简单

$$\xi = \sum_{i=1}^{n} r_i - 6. \tag{14.3.14}$$

利用以上两个公式产生的随机变量 ξ 的概率密度与 $N(0, 1)$ 的最大偏离出现在分布两侧尾端, 因为 $N(0, 1)$ 的取值域是 $(-\infty, \infty)$, 而 ξ 的取值域是有限值

$$-\sqrt{3n} \leqslant \xi \leqslant \sqrt{3n},$$

当 $n = 12$, 有 $-6 \leqslant \xi \leqslant 6$, 相当于 $N(0, 1)$ 的 ± 6 个标准差, 但是在那里, $N(0, 1)$ 概率密度值已小于 10^{-8}, 这表示 ξ 的概率密度值与标准正态的差别在 $\pm 6\sigma$ 处小于 10^{-8}. 对于大多数实际应用的场合, 这样小的误差是容许的.

由于 $N(\mu, \sigma^2)$ 的随机变量 Y 与 $N(0, 1)$ 随机变量 X 有简单的关系

$$Y = \mu + \sigma X,$$

根据式 (14.3.7) 立即可由 X 的随机数 ξ_i 求得 Y 的随机数 η_i

$$\eta_i = \mu + \sigma\xi_i. \tag{14.3.15}$$

同样, 德莫佛–拉普拉斯定理证明了正态分布是二项分布的极限情形, 即当 $n \to \infty$, 有

$$\frac{B(n,p) - np}{\sqrt{np(1-p)}} \sim N(0,1),$$

其中 $B(n, p)$ 为参数 n、p 的二项分布. 因此, 当 n 充分大, 二项分布的随机数 η_i 近似地可由下式求出 (式 (14.3.7)):

$$\eta_i = \xi_i\sqrt{np(1-p)} + np, \tag{14.3.16}$$

其中 ξ_i 是 $N(0, 1)$ 的随机数.

凡是存在极限分布的随机变量, 其随机数都可以类似的方式由极限分布的随机数求得.

14.3.5　复合分布的抽样方法

在实际问题中遇到的随机变量, 它们的概率密度常常不是简单的函数, 但有时可表示为若干个已知概率密度的求和, 或者乘积等等, 称为复合分布. 下面的几种抽样方法对于复合分布的抽样是很有用的.

(1) 加抽样方法

设随机变量的概率密度 $f(x)$ 可表示为若干个简单概率密度 $f_j(x)$ 之和

$$f(x) = \sum_{j=1}^{n} p_j f_j(x), \tag{14.3.17}$$

其中 $p_j \geqslant 0$, 为了保证 $f(x)$ 的归一性, 必须有

$$\sum_{j=1}^{n} p_j = 1.$$

令 η_j 为 $f_j(x)$ 的随机数, 则当

$$\sum_{j=1}^{m-1} p_j < r_i \leqslant \sum_{j=1}^{m} p_j \tag{14.3.18}$$

成立时, 有

$$\xi_i = \eta_m.$$

即 $f(x)$ 的随机数 ξ_i 取为 $f_m(x)$ 的随机数 η_m. 重复以上步骤, 即得 $f(x)$ 的随机数列 $\xi_1, \xi_2, \cdots$.

该方法的抽样效率显然取决于 $\eta_j(j=1,2,\cdots,n)$ 的抽样效率.

(2) 乘抽样方法

设随机变量的概率密度 $f(x)$ 可表示为

$$f(x) = H(x)g(x), \tag{14.3.19}$$

$g(x)$ 是任意概率密度函数, 其随机数令为 η_i. 由 $f(x)$ 和 $g(x)$ 的非负性知 $H(x)$ 为非负函数, 令 $H(x)$ 的极大值为 M. $f(x)$ 的随机数 ξ_i 抽样方法如图 14.5 所示. 与 14.3.3 节舍选抽样的框图对比可知, 本方法的抽样效率取决于两个因素：一个是 η_i 的抽样效率, 另一个是 $H(x)$ 的行为, $H(x)$ 变化较小则有较高的抽样效率.

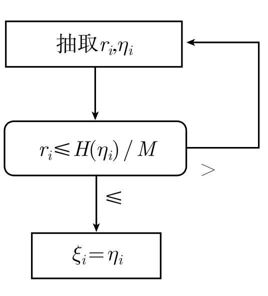


图 14.5　乘抽样方法

(3) 乘加抽样方法

设随机变量的概率密度 $f(x)$ 可表示为

$$f(x) = \sum_{j=1}^{n} H_j(x)g_j(x),$$

$g_j(x)(j=1,2,\cdots,n)$ 为任意概率密度函数, $H_j(x) \geqslant 0(j=1,2,\cdots,n)$.

不失一般性, 只考虑 $n=2$ 的情形

$$f(x) = H_1(x)g_1(x) + H_2(x)g_2(x). \tag{14.3.20}$$

将上式改写成

$$\begin{aligned} f(x) &= p_1 \frac{H_1(x)g_1(x)}{p_1} + p_2 \frac{H_2(x)g_2(x)}{p_2} \\ &\equiv p_1 g_1'(x) + p_2 g_2'(x), \end{aligned}$$

其中,

$$p_1 \equiv \int_{-\infty}^{\infty} H_1(x)g_1(x)\mathrm{d}x, \qquad p_2 \equiv \int_{-\infty}^{\infty} H_2(x)g_2(x)\mathrm{d}x. \tag{14.3.21}$$

由 $f(x)$ 的归一性, 即 $\displaystyle\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$ 可知, 必有

$$p_1 + p_2 = 1.$$

首先用加抽样方法确定 $f(x)$ 的随机数 ξ_i 应当取 $g_1'(x)$ 的随机数 η_1' 还是取 $g_2'(x)$ 的随机数 η_2', 然后用乘抽样方法决定 η_1'(或 η_2') 的值. 令 M_1、M_2 为

$H_1(x)$、$H_2(x)$ 的极大值; η_1、η_2 为 $g_1(x)$、$g_2(x)$ 的随机数, ξ_i 的抽样框图如图 14.6 所示.

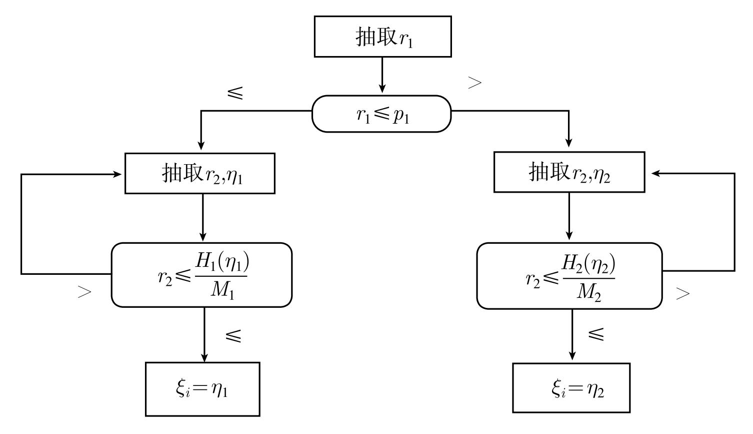


图 14.6　乘加抽样方法

14.3.6　近似抽样方法

实际问题中的概率密度的形式常常非常复杂, 有时甚至无法用解析函数的形式给出, 而只能用数值或曲线的形式表示. 在这种情形下, 就有必要用近似分布来代替原分布, 用近似分布的抽样作为原分布抽样的近似. 这种方法的优点首先在于它的普适性, 对任意的连续分布均适用. 其次, 只要子区间足够小, 分点足够多, 对于任意分布都可以达到相当好的近似程度. 同时方法也比较简便易行.

设原分布概率密度为 $f(x)(a \leqslant x \leqslant b)$, 将区间 $[a,b]$ 分成 n 个子区间, 分点为

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b, \tag{14.3.22}$$

分点对应的函数值为

$$f_i = f(x_i), \qquad i = 0, 1, 2, \cdots, n. \tag{14.3.23}$$

比较简单的方法是将 $[a,b]$ 区间等分, 但这样做近似分布与原分布差别较大. 较好的分法是使 $f_0, f_1, \cdots, f_n$ 能充分反映 $f(x)$ 的变化状况, 例如 $f(x)$ 变化迅速的区域分点可密一些, $f(x)$ 变化缓慢的区域分点可取得稀一些. 这样, 分点近似分布更接近于原分布.

(1) 阶梯近似

令

$$p_i = \int_{x_{i-1}}^{x_i} f(x)\mathrm{d}x. \tag{14.3.24}$$

利用阶梯函数 $f_a(x_i)$ 作为原分布概率密度的近似

$$f_a(x_i)=\int_{x_{i-1}}^{x_i}f(x)\mathrm{d}x/(x_i-x_{i-1}),\qquad i=1,2,\cdots,n. \tag{14.3.25}$$

每一子区间内原分布和近似分布的积分概率相同 (图 14.7). 显然

$$\sum_{i=1}^{n}p_i=\sum_{i=1}^{n}\int_{x_{i-1}}^{x_i}f(x)\mathrm{d}x=\int_{x_0}^{x_n}f(x)\mathrm{d}x=\int_a^b f(x)\mathrm{d}x=1.$$

$f_a(x_i)$ 的累积分布函数在分点 x_i 处的值为

$$F_a(x_i)=\sum_{j=1}^{i}p_j,\qquad i=1,2,\cdots,n. \tag{14.3.26}$$

近似分布 f_a 是每个子区间中的均匀分布, 因而其随机数 ξ_i 可如此选取：选取随机数 r_i, 找出满足

$$F_a(x_{j-1})<r_i\leqslant F_a(x_j) \tag{14.3.27}$$

的分点 x_{j-1} 和 x_j,ξ_i 可表示为

$$\xi_i=x_{j-1}+(x_j-x_{j-1})\frac{r_i-F_a(x_{j-1})}{F_a(x_j)-F_a(x_{j-1})}. \tag{14.3.28}$$

当 n 充分大, ξ_i 近似地可作为原分布 $f(x)$ 的随机数列.

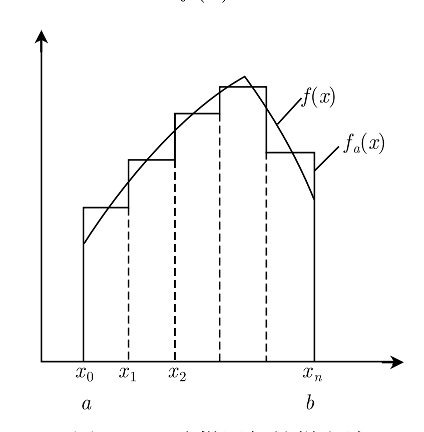


图 14.7 阶梯近似抽样方法

曲线是原分布 $f(x)$, 折线是阶梯近似分布 $f_a(x)$

(2) 线性近似

原分布的线性近似分布为

$$f_a(x)=C_i\left\{f_{i-1}+\frac{x-x_{i-1}}{x_i-x_{i-1}}(f_i-f_{i-1})\right\},$$

$$x_{i-1} < x \leqslant x_i, \qquad i = 1, 2, \cdots, n, \tag{14.3.29}$$

其中 C_i 为归一化因子, 使得每一子区间内原分布和近似分布的积分概率相同

$$\int_{x_{i-1}}^{x_i} f_a(x)\mathrm{d}x = \int_{x_{i-1}}^{x_i} f(x)\mathrm{d}x, \qquad i = 1, 2, \cdots, n.$$

这相当于原分布 $f(x)$ 连续函数用折线作为近似.

$f_a(x)$ 的随机数列 ξ_i 的抽样方法如图 14.8 所示.

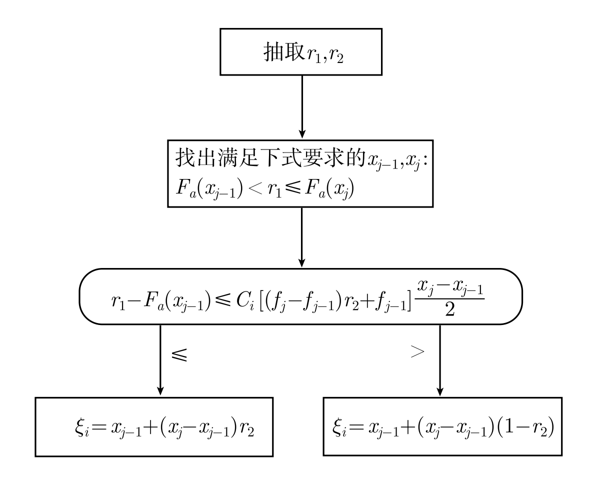


图 14.8 $f_a(x)$ 的随机数列 ξ_i 的抽样方法

当 n 充分大, ξ_i 近似地可作为 $f(x)$ 的随机数列.

比上述两种方法更精确的近似方法是二次曲线近似和样条近似, 即在各子区间内以二次曲线和样条函数作为原分布 $f(x)$ 的近似, 并保证在所有子区间内两者的积分概率相同. 其抽样方法的原则是相同的, 这里不再仔细介绍了.

14.3.7 多维分布的抽样

一维分布的许多随机抽样方法可推广运用于多维分布的随机抽样.

例如舍选法运用于多维抽样. 设 n 维概率密度为 $f(\boldsymbol{x})$, $\boldsymbol{x}$ 的取值域为 n 维长方体 $[\boldsymbol{a}, \boldsymbol{b}]$,

$$\boldsymbol{a} = \{a_1, a_2, \cdots, a_n\}, \qquad \boldsymbol{b} = \{b_1, b_2, \cdots, b_n\}.$$

$f(\boldsymbol{x})$ 的极大值为 M, 则 $f(\boldsymbol{x})$ 的随机数 $\boldsymbol{\xi} = \{\xi_1, \xi_2, \cdots, \xi_n\}$ 的舍选法如图 14.9 所示.

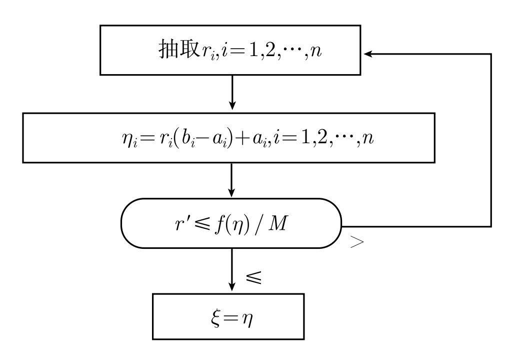


图 14.9 多维随机数 $\boldsymbol{\xi}$ 的舍选抽样

又如乘抽样方法应用于多维分布的抽样. 设多维随机变量概率密度函数 $f(\boldsymbol{x})$ 可表示为

$$f(\boldsymbol{x}) = H(\boldsymbol{x})g(\boldsymbol{x}),$$

其中 $g(\boldsymbol{x})$ 为任意多维概率密度函数, 其随机数为 $\boldsymbol{\eta} = \{\eta_1, \eta_2, \cdots, \eta_n\}$; $H(\boldsymbol{x}) \geqslant 0$ 的极大值令为 M, 则 $f(\boldsymbol{x})$ 的随机数 $\boldsymbol{\xi} = \{\xi_1, \xi_2, \cdots, \xi_n\}$ 有如图 14.10 所示的抽样方法.

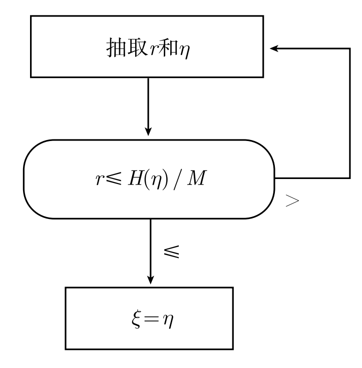


图 14.10 多维随机数 $\boldsymbol{\xi}$ 的乘抽样

14.3.6 节叙述的近似抽样也可以直接推广到多维分布的情形, 只需将一维子区间改变为 n 维子区间. 为了计算简单起见, 可以将子区间内的积分用近似值代替

$$\int_{\Delta x} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \approx f(\boldsymbol{x}^0)\Delta \boldsymbol{x},$$

其中 $\boldsymbol{x}^0$ 是子区间 $\Delta \boldsymbol{x}$ 内的某个点.

除了这些一维方法的推广之外, 多维分布的随机抽样还有所谓的条件密度法, 其实质是利用条件概率密度将多维模拟化为一维模拟问题. 任意 n 维概率密度 $f_n(x_1, \cdots, x_n)$ 可表示为

$$f_n(x_1, \cdots, x_n) = f(x_n|x_1, \cdots, x_{n-1}) \cdot f_{n-1}(x_1, \cdots, x_{n-1}). \tag{14.3.30}$$

右式第一项是 $x_1 = x_1, x_2 = x_2, \cdots, x_{n-1} = x_{n-1}$ 条件下的条件概率密度, 它是一维密度函数; 第二项

$$f_{n-1}(x_1, \cdots, x_{n-1}) = \int f_n(x_1, \cdots, x_n)\mathrm{d}x_n, \tag{14.3.31}$$

表 14.2　具有指定分布的随机数 η 的产生

分布	概率密度	变量域	随机数 η
$[a,b]$区间均匀分布	$\dfrac{1}{b-a}$	x: $[a,b]$	$r(b-a)+a$
	$\dfrac{1}{\lvert a\rvert}$	x: $\begin{matrix}[b-a,b] & (a>0)\\ [b+a,b] & (a<0)\end{matrix}$	$ar+b$　(b 任意实数)
	$\lambda\left(1-\dfrac{\lambda}{2}x\right)$　$(\lambda>0)$	$x:[0,2/\lambda]$	$\dfrac{2}{\lambda}\left(1-\sqrt{r}\right)$
	$\dfrac{c}{(1+bx)^2}$	$x:\left[0,\dfrac{1}{c-b}\right]$	$\dfrac{r}{c-br}$
	$nx^{n-1}(n\geqslant 1$ 整数$)$	$x:[0,1]$	$r^{1/n}$ 或 $\max(r_1,r_2,\cdots,r_n)$
	$\dfrac{1}{n}x^{\frac{1}{n}-1}(n\geqslant 1$ 整数$)$	$x:[0,1]$	r^n
	$nx^{-(n+1)}(n\geqslant 1$ 整数$)$	$x:[1,\infty)$	$1/\max(r_1,r_2,\cdots,r_n)$ 或 $r^{-1/n}$
	$\dfrac{2}{\pi\sqrt{1-x^2}}$	$x:[0,1]$	$\sin\pi r$
	$f(x)=\begin{cases}\dfrac{1}{2} & 0\leqslant x\leqslant 1\\ \dfrac{1}{2x^2} & 1<x<\infty\end{cases}$	$x:[0,\infty)$	$\dfrac{r_2}{r_1}$
指数分布	$\lambda e^{-\lambda x}$　$(\lambda>0)$	$x:[0,\infty)$	$-\dfrac{1}{\lambda}\ln r$
截尾指数分布	$\dfrac{\lambda e^{-\lambda x}}{1-e^{-\lambda x_m}}$　$(\lambda>0)$	$x:[0,x_m]$	$-\dfrac{1}{\lambda}\ln\left\{1-r\left[1-e^{-\lambda x_m}\right]\right\}$

续表

分布	概率密度	变量域	随机数 η
倒数分布	$\dfrac{1}{x\ln\lambda}\quad(\lambda>1)$	$x:[1,\lambda]$	λ^r
柯西分布	$\dfrac{1}{\pi(1+x^2)}$	$x:(-\infty,\infty)$	$\tan\pi(r-0.5)$
对数分布	$-(\ln x)^n/n!(n\geqslant 1$ 整数)	$x:(0,1]$	$\prod\limits_{i=1}^{n+1} r_i$
威布尔分布	$\dfrac{c}{b}\left(\dfrac{x-a}{b}\right)^{c-1}\exp\left\{-\left(\dfrac{x-a}{b}\right)^c\right\}$ $b,c>0,a\geqslant 0$	$x:[a,\infty)$	$a+b(-\ln r)^{1/c}$
圆内均匀分布	$2R/R_0^2$	$R:[0,R_0]$	$R_0\sqrt{r}$ 或 $R_0\max(r_1,r_2)$
环内均匀分布	$2R/\left(R_1^2-R_0^2\right)$	$R:[R_0,R_1]$	$\left[\left(R_1^2-R_0^2\right)r+R_0^2\right]^{1/2}$
球壳内均匀分布	$3R^2/\left(R_1^3-R_0^3\right)$	$R:[R_0,R_1]$	$\left[\left(R_1^3-R_0^3\right)r+R_0^3\right]^{1/3}$
n 维球内均匀分布		$x_i:[0,R_0]$ $i=1,2,\cdots,n$	$\left(\dfrac{R_0}{R}\xi_1,\dfrac{R_0}{R}\xi_2,\cdots,\dfrac{R_0}{R}\xi_n\right),R^2=\sum\limits_{i=1}^{n}\xi_i^2$
β 分布	$\dfrac{(N+1)!}{n!(N-n)!}x^n(1-x)^{N-n}\quad(N>n)$	$x:[0,1]$	产生 $N+1$ 个随机数 r_i, 且令 $r_1\leqslant r_2\leqslant\cdots\leqslant r_{N+1}$, 则 $\eta=r_{n+1}$
拉普拉斯分布	$\dfrac{1}{2}\exp\{-\|x\|\}$	$x:(-\infty,\infty)$	$\ln(r_1/r_2)$

续表

分布	概率密度	变量域	随机数 η
指数函数分布	$\int_1^{\infty} \mathrm{e}^{-xy}/y^n \mathrm{d}y$($n \geqslant 1$ 整数)	$x:[0,\infty)$	$-\max(r_1, r_2, \cdots, r_n)\ln r_{n+1}$
瑞利分布	$\frac{x}{\sigma^2}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}$	$x:[0,\infty)$	$\sigma\sqrt{-\ln r}$
	$\frac{a^n}{(n-1)!}x^{n-1}\mathrm{e}^{-a^x}$($n \geqslant 1$ 整数)	$x:[0,\infty)$	$-\frac{1}{a}\ln(r_1 \cdot r_2 \cdot \ \cdots \ \cdot r_n)$
半正态分布	$\sqrt{\frac{2}{\pi}}\exp\left(\frac{-x^2}{2}\right)$	$x:[0,\infty)$	$\sqrt{\frac{\pi}{8}}\ln\frac{1+r}{1-r}$, 随机产生正负号得到 ξ
两维正态分布	$\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}$	x, y $:(-\infty,\infty)$	$\eta_1 = \sqrt{-2\ln r_1}\cos 2\pi r_2$, $\eta_2 = \sqrt{-2\ln r_1}\sin 2\pi r_2$. η_1, η_2中任意一个, 都是一维 $N(0,1)$ 随机数 ξ
标准正态分布	$\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$	$x:(-\infty,\infty)$	$\sqrt{12n}\left(\frac{1}{n}\sum_{i=1}^{n} r_i - \frac{1}{2}\right)$, 常取 $n = 6, 12$.
			$y = \sqrt{-2\ln r}$(当满足$0.5 < r < 1$) $\eta = y - \frac{a_0 + a_1 y + a_2 y^2}{1 + b_1 y + b_2 y^2 + b_3 y^3}$, (误差 $< 10^{-4}$) $a_0 = 2.515517$　$a_1 = 0.802853$ $a_2 = 0.010328$ $b_1 = 1.432788$　$b_2 = 0.189269$ $b_3 = 0.001308$

续表

分布	概率密度	变量域	随机数 η
			$\eta = s\left[a_1 + s^2\left(a_3 + s^2[a_5 + s^2(a_7 + a_9 s^2)]\right)\right]$ $s = \left(\sum_{i=1}^{12} r_i - 6\right) \Big/ 4$(误差 $< 2 \times 10^{-4}$) $a_1 = 3.949846138 \quad a_3 = 0.252408784$ $a_5 = 0.076542912 \quad a_7 = 0.008355968$ $a_9 = 0.029899776$
正态分布	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$	$x:(-\infty,\infty)$	$\xi\sigma + \mu$
对数正态分布	$\frac{1}{\sqrt{2\pi}\sigma(x-a)} \exp\left\{\frac{[\ln(x-a)+A]^2}{-2\sigma^2}\right\}$	$x:(a,\infty)$	$a + \exp(\sigma\xi - A)$
$\chi^2(n)$ 分布		$x:[0,\infty)$	$\delta_n = \sum_{i=1}^{n} \xi_i^2.$ $\delta_{n=2k} = -2\ln(r_1 \cdot r_2 \cdot \ \cdots \ \cdot r_k)$ $\delta_{n=2k+1} = -2\ln(r_1 \cdot r_2 \cdot \ \cdots \ \cdot r_k) + \xi^2$
$\chi^2(n_1+n_2)$ 分布		$x:[0,\infty)$	$\delta_{n1} + \delta_{n2}$
$t(n)$ 分布		$x:(-\infty,\infty)$	$\xi \Big/ \sqrt{\frac{\delta_n}{n}}$
$F(m,n)$ 分布		$x:[0,\infty)$	$\frac{\delta_m/m}{\delta_n/n}$

表中 r, r_i——[0, 1] 区间均匀分布随机数; δ_n——$\chi^2(n)$ 分布随机数; ξ, ξ_i—— 标准正态分布 $N(0,\ 1)$ 随机数

它等于 $f_n(x_1,\cdots,x_n)$ 对其中一个变量 x_n 作积分, 是余下的 $n-1$ 个变量的联合概率密度. 反复运用式 (14.3.30), 可得

$$\begin{aligned}f_n(x_1,\cdots,x_n)=&f(x_n\,|x_1,\cdots,x_{n-1}\,)\\&\cdot f(x_{n-1}\,|x_1,\cdots,x_{n-2}\,)\cdots f(x_2\,|x_1\,)f(x_1).\end{aligned}\qquad(14.3.32)$$

该式右侧每项都是一维密度函数, 其中条件概率密度可从式 (14.3.30) 的变形求得

$$f(x_n\,|x_1,\cdots,x_{n-1}\,)=\frac{f_n(x_1,\cdots,x_n)}{f_{n-1}(x_1,\cdots,x_{n-1})}.\qquad(14.3.33)$$

由式 (14.3.32) 可得如下模拟方法: 先选出一维概率密度 $f_1(x_1)$ 的随机数 ξ_1, 并代入 $f(x_2\,|x_1\,)$ 中的 x_1 得到确定的一维概率密度 $f(x_2\,|\xi_1\,)$. 再选出分布 $f(x_2\,|\xi_1\,)$ 的随机数 ξ_2. 照此方式进行下去即可得到分布 $f_n(x_1,\cdots,x_n)$ 的随机数向量 $(\xi_1,\cdots,\xi_n)$. 由其抽样过程可知多维分布抽样的条件密度法要求对多维分布有相当完整的知识, 即必须知道边沿分布 $f_1(x_1)$ 和条件分布 $f(x_2\,|x_1\,), f(x_3\,|x_1,x_2\,),\cdots,f(x_n|x_1,\cdots,x_{n-1})$ 的函数形式.

表 14.2 我们列出了一些分布的随机数产生的方法. 多数是用直接抽样方法得到的, 也有一些使用了其他抽样技巧. 表中列出的随机数产生方法不是唯一的, 抽样费用也不一定最省.

14.4 蒙特卡罗法计算积分

计算积分和多重积分是蒙特卡罗法的重要应用领域之一. 对于积分边界复杂、被积函数形式复杂的积分, 用解析方法往往难以求解, 用一般的数值方法也感到困难, 蒙特卡罗法总能比较简单地求出其近似解及其误差.

14.4.1 频率法 (均匀投点法)

在 14.1 节中已经介绍过均匀投点法求最简单的一维积分的例子, 现在讨论

$$I=\int_a^b g(x)\mathrm{d}x.\qquad(14.4.1)$$

$g(x)$ 是 $[a,b]$ 区间上非负、有界可积函数, 其极大值为 M. 积分值就是图 14.11 中划斜线的面积 A. 抽取 n 对随机数

$$\{\xi_i,\eta_i\},\qquad i=1,2,\cdots,n,$$

其中 ξ_i 是 $[a,b]$ 区间均匀分布的随机数, η_i 是 $[0,M]$ 区间均匀分布的随机数

$$\xi_i=a+(b-a)r_i,\qquad \eta_i=Mr_j,\qquad(14.4.2)$$

其中 r_j 是不同于 r_i 的 [0,1] 区间均匀分布的随机数. 若满足

$$g(\xi_i) > \eta_i, \tag{14.4.3}$$

则点 $\{\xi_i, \eta_i\}$ 落在面积 A 内, 设共有 m 个点 (即 m 对 ξ_i, η_i) 落在面积 A 内, 全部点数为 n, 则不难证明, 当 n 充分大时有下列近似公式:

$$I = \int_a^b g(x)\mathrm{d}x \approx \frac{m}{n} M(b-a) \equiv I_n. \tag{14.4.4}$$

因而可算得积分 I 的近似值 I_n.

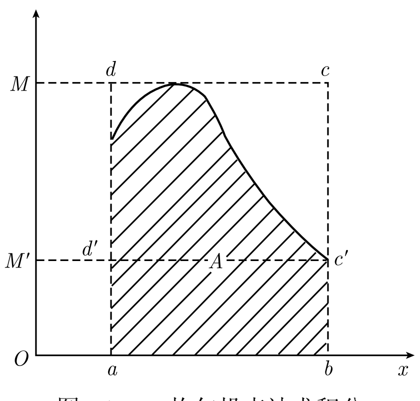


图 14.11　均匀投点法求积分

事实上, 该算法相当于在矩形 $a \leqslant x \leqslant b, 0 \leqslant y \leqslant M$ 中均匀、独立地投点, 假定点落在面积 A 中为成功, 则每次投点成功的概率 p 等于面积 A 与矩形 $abcd$ 面积之比

$$p = \frac{A}{(b-a)M}.$$

同时, 由大数定律知

$$\lim_{n\to\infty} \frac{m}{n} = p.$$

故当 n 足够大时, 式 (14.4.4) 成立.

n 次投点可看成 n 次伯努利试验, 成功的概率为 p, 所以成功的次数 m 服从参数 n, p 的二项分布, 其均值和方差分别为

$$E(m) = np, \qquad \sigma^2(m) = np(1-p).$$

从而有

$$\begin{aligned} E(I_n) &= E\left[m \cdot \frac{M(b-a)}{n}\right] = \frac{M(b-a)}{n} E(m) \\ &= M(b-a)p = A = I, \end{aligned}$$

$$\sigma^2(I_n)=\sigma^2\left[m\cdot\frac{M(b-a)}{n}\right]=\frac{M^2(b-a)^2}{n^2}\sigma^2(m)$$
$$=M^2(b-a)^2p(1-p)/n. \tag{14.4.5}$$

可见, I_n 是积分值 I 的无偏估计量, 而积分估计值的误差是

$$\sigma(I_n)=M(b-a)\sqrt{\frac{p(1-p)}{n}}. \tag{14.4.6}$$

因为 p 的真值是不知道的, 当 n 充分大时, $p\approx m/n$, 于是有

$$\sigma(I_n)\approx\frac{M(b-a)}{n}\sqrt{m\left(1-\frac{m}{n}\right)}. \tag{14.4.7}$$

一个重要的问题是, 对于任意的给定正数 ε, 怎样才能保证积分计算值 I_n 与真值 I 之差的绝对值 Δ 小于 ε 的概率不小于 $\alpha(0<\alpha<1)$. Δ 可表示为

$$\Delta=|I_n-I|=M(b-a)\left|\frac{m}{n}-p\right|.$$

根据中心极限定理, 当 n 充分大, 随机变量

$$\frac{m-np}{\sqrt{np(1-p)}}\sim N(0,1).$$

因此, 有

$$P(\Delta<\varepsilon)=P\left\{M(b-a)\left|\frac{m}{n}-p\right|<\varepsilon\right\}=P\left\{\frac{|m-np|}{\sqrt{np(1-p)}}<t\right\}$$
$$=\varPhi(t)-\varPhi(-t)=2\varPhi(t)-1,$$

其中,

$$t=\frac{\varepsilon}{M(b-a)}\sqrt{\frac{n}{p(1-p)}}. \tag{14.4.8}$$

于是问题化为求出适当的 t 值, 满足

$$\alpha\leqslant P(\Delta<\varepsilon)=2\varPhi(t)-1.$$

对于给定的 α, 可从累积标准正态分布表查出满足

$$\alpha=2\varPhi(t_\alpha)-1 \tag{14.4.9}$$

的 t_α 值. 令 $t=t_\alpha$, 由式 (14.4.8) 求得

$$\varepsilon_\alpha=t_\alpha M(b-a)\sqrt{\frac{p(1-p)}{n}}. \tag{14.4.10}$$

这表示在置信水平 α 上, 积分计算值与真值之差的上限为 ε_α, 它与投点数 n 的根号成反比. 投点数越多, 误差越小. 与式 (14.4.6) 对比, 上式也可写成

$$\varepsilon_\alpha = t_\alpha \cdot \sigma(I_n). \tag{14.4.11}$$

反过来, 为了在置信水平 α 上, 积分计算值与真值之差小于给定正数 ε, 应有 $t \geqslant t_\alpha$, 即投点数 n 须满足

$$n \geqslant \frac{t_\alpha^2 M^2 (b-a)^2 p(1-p)}{\varepsilon_\alpha^2}. \tag{14.4.12}$$

其中 p 是未知的, 当 n 充分大时可用 m/n 作为 p 的近似值.

上述计算方法的基本思想在 14.1 节的例子中已经用到过. 由于积分是利用随机变量在一定区域内的频率来计算的, 概率模型是一定区域内的均匀投点, 故称为均匀投点法或频率法.

我们还可以利用被积函数的性质减小积分估计值的误差. 如果被积函数 $g(x)$ 在积分区域 $[a,b]$ 内存在大于 0 的极小值 M', 那么图 14.11 中 $abc'd'$ 的面积十分容易计算, 无须运用蒙特卡罗技巧. 积分可分为两部分

$$\begin{aligned} I &= I' + \int_a^b [g(x) - M']\,\mathrm{d}x \equiv I' + I'', \\ I' &= M'(b-a). \end{aligned} \tag{14.4.13}$$

其中 I'' 可用以上介绍的方法求积分. 这时前面求得的结果直接可用, 唯一的更改是将原式中的 M 改为 $M - M'$. 由式 (14.4.6) 知道, 积分的误差 $\sigma(I_n'')$ 比 $\sigma(I_n)$ 小.

均匀投点法原则上也适用于多重积分. 例如, 求多重积分

$$I = \int_a^b g(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_l}^{b_l} g(x_1, x_2, \cdots, x_l)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_l,$$

令 $\boldsymbol{\xi} = \{\xi_1, \xi_2, \cdots, \xi_l\}$ 是 $[\boldsymbol{a}, \boldsymbol{b}]$ 区间内均匀分布的随机数, $\eta = Mr$ 是 $[0, M]$ 区间内均匀分布随机数, M 是 $g(\boldsymbol{x})$ 在 $[\boldsymbol{a}, \boldsymbol{b}]$ 区间内的极大值. 选取 n 组 $\{\boldsymbol{\xi}, \eta\}$, 若其中满足

$$g(\boldsymbol{\xi}) > \eta$$

的共有 m 组, 则当 n 充分大时,

$$I = \int_a^b g(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \approx \frac{m}{n} M \prod_{i=1}^{l} (b_i - a_i) \equiv I_n. \tag{14.4.14}$$

I_n 可作为积分的近似值. 关于 I_n 的方差, I_n 与积分真值之差小于给定正数 ε 所要求的 n 值, 可用类似于一维的方式进行讨论. 只要将一维公式中的 $(b-a)$ 改为 $\prod\limits_{i=1}^{l}(b_i - a_i)$ 便可应用于多维积分的场合.

实验工作者关心的通常是相对误差, 即积分值的误差与积分值之比. 由式 (14.4.5) 和式 (14.4.10) 知, 在给定置信水平 α 上, 积分的相对误差 δ 可表示为

$$\delta = t_\alpha \sqrt{\frac{1-p}{pn}} = \frac{t_\alpha}{\sqrt{n}} \left(\frac{1}{p} - 1\right)^{1/2}. \tag{14.4.15}$$

该公式对于一维或多维积分都正确. 它说明利用均匀投点法求积分的精度与两个因素有关. 第一是与投点数 n 的平方根成反比, n 越大, 精度越好, 而且与积分的维数无关. 第二, 精度与被积函数 $g(x)$ 的行为有关, $g(x)$ 变化越平缓, 投点成功的概率 p 越接近于 1, 精度越好. 一般说来, 如果对于函数的行为无所了解, 尤其在高维积分的情形下, 投点成功的概率 p 往往很小, 这样精度就很差, 所以均匀投点法一般地不适合求高维积分.

14.4.2 期望值估计法

任何一个积分都可以表示为某个随机变量的数学期望, 因此, 可以用该随机变量的子样平均值作为积分的近似值.

设欲求的积分为

$$I = \int_{V_s} g(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}, \tag{14.4.16}$$

其中 $\boldsymbol{x} = \{x_1, x_2, \cdots, x_s\}$ 表示 S 维空间的点, V_s 表示积分区域. 令 $f(\boldsymbol{x})$ 为 V_s 上的任一随机变量 $\boldsymbol{\xi}$ 的概率密度函数

$$\int_{V_s} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = 1, \tag{14.4.17}$$

则积分 I 可表示为随机变量 $h(\boldsymbol{x}) = g(\boldsymbol{x})/f(\boldsymbol{x})$ 的数学期望

$$I = \int_{V_s} g(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \int_{V_s} \frac{g(\boldsymbol{x})}{f(\boldsymbol{x})} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = E\left[\frac{g(\boldsymbol{x})}{f(\boldsymbol{x})}\right] = E[h(\boldsymbol{x})]. \tag{14.4.18}$$

当从随机变量 $\boldsymbol{\xi}$ 抽取容量为 n 的随机子样 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_n$(即服从分布 $f(\boldsymbol{x})$ 的随机数), 可求得随机变量 $h(\boldsymbol{x})$ 的子样 $h_1, h_2, \cdots, h_n$,

$$h_i = h(\boldsymbol{\xi}_i) = g(\boldsymbol{\xi}_i) / f(\boldsymbol{\xi}_i). \tag{14.4.19}$$

而根据大数定律, 当 $n \to \infty$ 时, 子样平均

$$\hat{h} = \frac{1}{n} \sum_{i=1}^{n} h_i \tag{14.4.20}$$

的期望值 $E[\hat{h}]$ 与总体 h 的数学期望 $E[h]$ 相等, 所以当 n 充分大时, 有

$$I = E[h] = E[\hat{h}] \approx \frac{1}{n} \sum_{i=1}^{n} h_i \equiv I_n. \tag{14.4.21}$$

选取 $f(\boldsymbol{x})$ 的最简单方法是 V_s 内的均匀分布

$$f(\boldsymbol{x}) = 1/V_s,$$

这样, 上式简化为

$$I \approx I_n \equiv \frac{V_s}{n}\sum_{i=1}^{n} g\left(\boldsymbol{\xi}_i\right), \tag{14.4.22}$$

这里, $\boldsymbol{\xi}_i$ 是 V_s 中均匀分布的随机数列, 它们很容易从 $x_1, x_2, \cdots, x_s$ 各自的积分限和随机数 r 求出. 对于一维积分的特殊情形, 若积分上下限为 b 和 a, 则立即有

$$I \approx I_n = \frac{(b-a)}{n}\sum_{i=1}^{n} g(\xi_i). \tag{14.4.23}$$

容易求得式 (14.4.21) 积分近似值 I_n 的期望值和方差

$$E(I_n) = E\left[\frac{1}{n}\sum_{i=1}^{n} h_i\right] = E(h) = I,$$

$$V(I_n) = V\left[\frac{1}{n}\sum_{i=1}^{n} h_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} V\left(h_i\right) = \frac{1}{n}V_h,$$

其中 V_h 是随机变量 $h(\boldsymbol{x})$ 的方差

$$V_h = \int_{V_s}[h-E(h)]^2 f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_{V_s}(h-I)^2 f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}. \tag{14.4.24}$$

根据同分布中心极限定理, 当 $n \to \infty$ 时, 有

$$\frac{I_n - I}{\sqrt{V_h/n}} \sim N(0,1).$$

利用这一极限性质可以得到积分估计值 I_n 与真值 I 之差 $\Delta = |I_n - I| < \varepsilon$($\varepsilon$ 给定正数) 的概率不小于 $\alpha(0<\alpha<1)$ 所要求的条件. 这时有

$$\begin{aligned}\alpha \leqslant P\{|I_n - I| < \varepsilon\} &= P\left\{\frac{|I_n - I|}{\sqrt{V_h/n}} < \frac{\varepsilon}{\sqrt{V_h/n}}\right\}\\ &= \varPhi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{V_h}}\right) - \varPhi\left(\frac{-\varepsilon\sqrt{n}}{\sqrt{V_h}}\right) = 2\varPhi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{V_h}}\right) - 1.\end{aligned}$$

对于给定的 α, 可从累积标准正态分布表查出满足

$$\alpha = 2\varPhi(t_\alpha) - 1 \tag{14.4.25}$$

的 t_α 值, 令

$$t_\alpha = \frac{\varepsilon_\alpha \sqrt{n}}{\sqrt{V_h}},$$

求得

$$\varepsilon_\alpha = \frac{t_\alpha \sqrt{V_h}}{\sqrt{n}}. \tag{14.4.26}$$

这表示当子样容量为 n 时, 在置信水平 α 上, 积分估计值 I_n 与真值 I 之差 $\varDelta$ 的上限为 ε_α.

反过来, 为了在置信水平 α 上有 $\varDelta < \varepsilon$, 应有 $t \geqslant t_\alpha$, 因而求出

$$n \geqslant t_\alpha^2 V_h / \varepsilon^2. \tag{14.4.27}$$

由该式可算得对于给定误差 ε 所要求的子样容量 n.

问题是上面两个公式中的 V_h 是未知量, 但由式 (14.4.24) 有

$$\begin{aligned} V_h &= \int_{V_s} (h-I)^2 f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \int_{V_s} h^2 f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} - I^2 = E(h^2) - I^2 \\ &\approx \frac{1}{n} \sum_{i=1}^{n} h^2 (\xi_i) - \left[\frac{1}{n} \sum_{i=1}^{n} h(\xi_i) \right]^2 \equiv \hat{h}^2 - \left(\hat{h}\right)^2. \end{aligned}$$

因此, 积分近似值的方差为

$$V(I_n) = \frac{1}{n} V_h \approx \frac{1}{n} \left[\hat{h}^2 - \left(\hat{h}\right)^2 \right]. \tag{14.4.28}$$

为了使积分的估计值 I_n 达到某个预定的精度 ε, 可采用以下步骤.

设已进行了 n 次抽样, 对于给定的置信水平 α, 利用式 (14.4.26) 和式 (14.4.28) 求出此时的误差 ε_n. 若 $\varepsilon_n < \varepsilon$, 则预定精度已经达到, $I_n = \hat{h}_n$ 即为所需的积分值; 若 $\varepsilon_n > \varepsilon$, 则需进一步增加抽样次数.

显然, 当 n 很大时, 利用式 (14.4.21) 和式 (14.4.28) 计算积分 $I_n = \hat{h}_n$ 及其误差 $V(I_n)$ 是比较费时的, 当 $h(\boldsymbol{x})$ 形式复杂时尤其如此. 因此, 当抽样次数由 n 增加到 $n+1$, 计算 $\hat{h}_{n+1}$ 和 $V(I_{n+1})$ 时若能利用 $\hat{h}_n$ 和 $V(I_n)$ 的值将能大大节省计算机的机时. 下面来推导计算 $\hat{h}_{n+1}$ 和 $V(I_{n+1})$ 的递推公式.

对于积分值 $I_{n+1} = \hat{h}_{n+1}$, 有

$$\hat{h}_{n+1} = \frac{1}{n+1} \sum_{i=1}^{n+1} h(\boldsymbol{\xi}_i) = \frac{1}{n+1} \left[\sum_{i=1}^{n} h(\boldsymbol{\xi}_i) + h(\boldsymbol{\xi}_{n+1}) \right],$$

故得

$$I_{n+1} \equiv \hat{h}_{n+1} = \frac{1}{n+1} \left[n\hat{h}_n + h(\boldsymbol{\xi}_{n+1}) \right]. \tag{14.4.29}$$

而 I_{n+1} 的方差为 (见式 (14.4.28))

$$V(I_{n+1})=\frac{1}{n+1}\left[\hat{h}_{n+1}^2-\left(\hat{h}_{n+1}\right)^2\right],$$

其中方括号中第一项

$$\begin{aligned}\hat{h}_{n+1}^2&=\frac{1}{n+1}\sum_{i=1}^{n+1}h^2\left(\boldsymbol{\xi}_i\right)=\frac{1}{n+1}\left[h^2(\boldsymbol{\xi}_{n+1})+\sum_{i=1}^{n}h^2\left(\boldsymbol{\xi}_i\right)\right]\\&=\frac{1}{n+1}\left[h^2(\boldsymbol{\xi}_{n+1})+n\hat{h}_n^2\right]\\&=\frac{1}{n+1}\left[h^2(\boldsymbol{\xi}_{n+1})+n^2V(I_n)+n(\hat{h}_n)^2\right],\end{aligned}$$

所以有

$$\begin{aligned}V(I_{n+1})=&\frac{1}{(n+1)^2}[h^2(\boldsymbol{\xi}_{n+1})+n^2V(I_n)\\&+n(\hat{h}_n)^2-(n+1)(\hat{h}_{n+1})^2].\end{aligned}\tag{14.4.30}$$

这样, 就可由新产生的随机数 $\boldsymbol{\xi}_{n+1}$ 对应的函数值 $h(\boldsymbol{\xi}_{n+1}),h^2(\boldsymbol{\xi}_{n+1})$ 以及已有的 $\hat{h}_n, V(I_n)$ 求得第 $n+1$ 次抽样后的积分估计值 $I_{n+1}=\hat{h}_{n+1}$ 及其方差 $V(I_{n+1})$.

14.4.3　重要抽样方法

由式 (14.4.26) 可知, 在积分的期望值估计法中, 对于一定的置信水平, 积分误差的上界 ε_α 取决于函数 $h(\boldsymbol{x})$ 的方差 V_h 和抽样次数 $\sqrt{n}$. 相应地可以通过增大抽样次数 n 和减小方差 V_h 这两种途径来减小积分的误差. 重要抽样是降低方差 V_h 的一种常用方法.

随机变量 $h(\boldsymbol{x})$ 的方差为

$$\begin{aligned}V_h&=\int_{V_s}(h-I)^2f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}=E(h^2)-[E(I)]^2\\&=\int_{V_s}\frac{g^2(\boldsymbol{x})}{f^2(\boldsymbol{x})}f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}-I^2.\end{aligned}\tag{14.4.31}$$

现在的问题是怎样选择适当的概率密度 $f(\boldsymbol{x})$, 使得 V_h 达到极小.

可以证明, 如果选择 $f(\boldsymbol{x})$ 为

$$f(\boldsymbol{x})=\frac{|g(\boldsymbol{x})|}{\displaystyle\int_{V_s}|g(\boldsymbol{x})|\mathrm{d}\boldsymbol{x}}\tag{14.4.32}$$

V_h 达到极小值. 特别当 $g(\boldsymbol{x})\geqslant 0$ 时上式相当于

$$f(\boldsymbol{x})=\frac{g(\boldsymbol{x})}{\displaystyle\int_{V_s} g(\boldsymbol{x})\mathrm{d}\boldsymbol{x}}=\frac{g(\boldsymbol{x})}{I}. \tag{14.4.33}$$

代入 V_h 的表达式, 得

$$V_h=\int_{V_s}\frac{g^2(\boldsymbol{x})}{g^2(\boldsymbol{x})}I^2 f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}-I^2=I^2-I^2=0,$$

即方差为零. 这种使 V_h 方差等于零的技巧称为重要抽样.

但是从式 (14.4.32) 和式 (14.4.33) 可见, 为了能构成重要抽样这样的概率密度函数 $f(\boldsymbol{x})$, 首先要求出积分值或同等工作量的积分

$$\int_{V_s}|g(\boldsymbol{x})|\mathrm{d}\boldsymbol{x}.$$

因此, 零方差重要抽样实际上不可能实行. 但是重要抽样的原则提供了寻找较优的 $f(\boldsymbol{x})$ 的途径. 从式 (14.4.33) 可知, 对于零方差情形,

$$f(\boldsymbol{x})\propto g(\boldsymbol{x}), \tag{14.4.34}$$

这相当于在 $g(\boldsymbol{x})$ 大的地方 (对积分 I 贡献大的区域), $f(\boldsymbol{x})$ 也选得较大, 即抽样点取得密一些; 而在 $g(\boldsymbol{x})$ 小的地方, 抽样点取得稀一些. 因此, 当概率密度函数 $f(\boldsymbol{x})$ 取得近似地正比于被积函数 $g(\boldsymbol{x})$ 时再用期望值估计法估计积分值 I, 便能大大减小方差 V_h, 从而大大减小积分值的方差 $V(I_n)=V_h/n$.

14.4.4 半解析法

这种方法把解析计算方法和概率计算方法结合起来, 将待求积分的主要部分用没有误差的解析方法求积, 因而大大降低了积分的误差.

设待求积分为

$$I=\int_{V_s} g(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

如能找到一个函数

$$g'(\boldsymbol{x})\approx g(\boldsymbol{x}) \tag{14.4.35}$$

而积分

$$J=\int_{V_s} g'(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{14.4.36}$$

为已知或可解析求解, 则有

$$I=J+\int_{V_s}[g(\boldsymbol{x})-g'(\boldsymbol{x})]\mathrm{d}\boldsymbol{x}.$$

该式表明, 只需要对 $\displaystyle\int_{V_s}[g(\boldsymbol{x})-g'(\boldsymbol{x})]\mathrm{d}\boldsymbol{x}$ 用蒙特卡罗法求解. 令

$$h = \frac{g(\boldsymbol{x}) - g'(\boldsymbol{x})}{f(\boldsymbol{x})} + J, \tag{14.4.37}$$

则随机变量 h 的期望值为

$$\begin{aligned} E(h) &= \int_{V_s} hf(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_{V_s} [g(\boldsymbol{x}) - g'(\boldsymbol{x})]\mathrm{d}\boldsymbol{x} + J \\ &= I - J + J = I. \end{aligned} \tag{14.4.38}$$

即 $E(h)$ 等于积分值. 因此, 当 n 充分大时, 随机变量 h 的子样平均 $\hat{h}$ 可作为积分 I 的近似值

$$I \approx \hat{h} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{g(\boldsymbol{\xi}_i) - g'(\boldsymbol{\xi}_i)}{f(\boldsymbol{\xi}_i)}\right] + J, \tag{14.4.39}$$

其中 $\boldsymbol{\xi}_i$ 为 $f(\boldsymbol{x})$ 的随机数.

随机变量 h 的方差这时为

$$\begin{aligned} V_h &= \int_{V_s} (h - I)^2 f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ &= \int_{V_s} \left[\frac{g(\boldsymbol{x}) - g'(\boldsymbol{x})}{f(\boldsymbol{x})} - I + J\right]^2 f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\ &= \int_{V_s} \left[\frac{g(\boldsymbol{x}) - g'(\boldsymbol{x})}{f(\boldsymbol{x})}\right]^2 f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - (I - J)^2. \end{aligned} \tag{14.4.40}$$

由于 $g(\boldsymbol{x}) \approx g'(\boldsymbol{x}), I \approx J$, 所以上式右边的两项比 V_h 的一般表达式 (14.4.31) 中相对应的两项要小得多, 从而减小了积分估计值的误差.

例 14.1 蒙特卡罗法计算积分 $I = \int_0^1 \mathrm{e}^x \mathrm{d}x$

积分 $I = \int_0^1 \mathrm{e}^x \mathrm{d}x$ 是解析可积的, 积分值为 $I = 1.718\,28$. 我们现在用 14.4.1 节~14.4.4 节介绍的四种方法求 I 的近似值, 以比较它们的精度. 子样容量统一地取为 $n = 10$.

对于投点法, 立即得到积分近似值公式如下 (式 (14.4.4)):

$$I_1 = \frac{\mathrm{e}}{n}\sum_{i=1}^{n} g(r_i, \eta_i),$$

其中,

$$g(r_i, \eta_i) = \begin{cases} 1, & 若\mathrm{e}^{r_i} > \eta_i, \\ 0, & 若\mathrm{e}^{r_i} \leqslant \eta_i. \end{cases}$$

r_i 是 [0, 1] 区间均匀分布随机数, η_i 是 $[0, e]$ 区间内均匀分布随机数.

当用期望值估计法求积分, 有 (式 (14.4.22))

$$I_2 = \frac{1}{n}\sum_{i=1}^{n} \mathrm{e}^{r_i}.$$

利用重要抽样方法求积分时需要选择适当的概率密度函数 $f(x)$. 因为

$$\mathrm{e}^x = 1 + x + \cdots ,$$

所以选择

$$f(x) = \frac{2}{3}(1+x),$$

其中的乘因子是为了保证 $f(x)$ 在 [0, 1] 积分区间内的归一性. $f(x)$ 的随机数 ξ_i 可由反函数法求出

$$r_i = \int_0^{\xi_i} f(x)\mathrm{d}x = \frac{2}{3}\left(\xi_i + \frac{\xi_i^2}{2}\right),$$

解得

$$\xi_i = -1 \pm \sqrt{1+3r_i}.$$

由于 x 的取值在 [0, 1] 区间内, ξ_i 的两个解中只有

$$\xi_i = \sqrt{1+3r_i} - 1$$

符合要求. 计算积分的公式于是为

$$I_3 = \frac{1}{n}\sum_{i=1}^{n} \frac{\mathrm{e}^{\xi_i}}{f(\xi_i)} = \frac{3}{2n}\sum_{i=1}^{n} \frac{\mathrm{e}^{\xi_i}}{(1+\xi_i)}.$$

最后给出用半解析法计算该积分的公式. 取 $g'(x) = 1+x, f(x) = 1$(即 [0, 1] 区间内均匀分布), 代入式 (14.4.39), 得

$$\begin{aligned} I_4 &= \frac{1}{n}\sum_{i=1}^{n} (\mathrm{e}^{r_i} - 1 - r_i) + \int_0^1 (1+x)\mathrm{d}x \\ &= \frac{1}{n}\sum_{i=1}^{n} (\mathrm{e}^{r_i} - r_i) + \frac{1}{2}. \end{aligned}$$

利用以下 10 个随机数 r_i:

$$0.865\,15,\ 0.907\,95,\ 0.661\,55,\ 0.664\,34,\ 0.565\,58,$$

$$0.123\,32,\ 0.943\,77,\ 0.578\,02,\ 0.691\,86,\ 0.033\,93,$$

代入 I_1、I_2、I_3、I_4 的表达式 ($n = 10$), 求得表 14.3 所列的积分近似值, 表中还列出了积分的误差 σ.

表 14.3 四种蒙特卡罗求积分法的比较

计算方法	频率法	期望值法	重要抽样	半解析法
积分近似值 I'	1.359	1.901	1.782	1.798
$I'-I$	−0.359	0.183	0.064	0.080
$V(I')$	0.1718	0.0242	0.00270	0.00437
$\sigma(I')$	0.414	0.156	0.052	0.066

从表 14.3 中数值可以看到, 重要抽样和半解析法有较好的精度. 虽然抽取的子样容量 ($n=10$) 不够大, 但求得的误差 σ 大致上反映了积分近似值与真值的差别.

14.4.5 自适应蒙特卡罗积分

对于式 (14.4.16) 所示的积分

$$I=\int_{V_s} g(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

如果被积函数 $g(\boldsymbol{x})$ 的形式极为复杂, 我们对于函数在被积区域 V_s 中的行为基本上无所了解, 则不能运用 14.4.3 节和 14.4.4 节所叙述的方法. 虽然仍可用频率法或期望值估计法求积分的近似值, 但方差可能很大.

此外, 如果被积函数在积分边界上有 (可积) 奇点, 前面叙述的几种方法不能使用. 但是在量子电动力学的计算中经常遇到这种类型的积分.

在出现上述情形时, 可应用自适应蒙特卡罗求积分的技巧求出积分近似值及其误差. 它的基本步骤如下: 在积分区域 V_s 内将 s 根轴划分为若干长度元, 于是 V_s 被划分为一组 s 维体积元, 在每个体积元中通过前面介绍的任一种方法 (如期望值估计法) 求出被积函数 $g(\boldsymbol{x})$ 的积分值 I_i 及其方差 σ_i^2. 利用 I_i 和 σ_i^2 可重新对 s 根轴作更合理的划分, 得到一组新的 s 维体积元, 在这组体积元中求出的积分值具有较小的方差. 如此迭代下去, 直到计算得到的积分值的误差小于预先给定的限值便中止迭代.

设积分区域 V_s 被划分为 N 个体积元 V_i,

$$V_s=\sum_{i=1}^{N} V_i.$$

在每个体积元中利用期望值方法计算积分, 并选择概率密度函数 $f(\boldsymbol{x})$ 为 V_i 中的均匀分布, 每个体积元中的抽样点数为 L_i, 因此, 整个积分体积 V_s 中总的抽样点数 L 为

$$L=\sum_{i=1}^{N} L_i,$$

这样, 在 V_i 中抽样点的密度可表示为

$$P_i = \frac{1}{L}\frac{L_i}{V_i}, \qquad i = 1, 2, \cdots, N. \tag{14.4.41}$$

P_i 的这种表示具有归一性, 即

$$\sum_{i=1}^{N} P_i V_i = 1.$$

在重要抽样中已经阐明, 为了减小积分的误差, 在函数值 $g(\boldsymbol{x})$ 大的区域抽样点密度应当大一些, 反之则密度应小一些. 所以抽样点密度 P_i 应按这样的原则来选择. 但在第一次迭代中, 因为对函数的行为无所了解, 所以选择在每个体积元中有相同的密度 P_i

$$P_i = \frac{1}{V_s},$$

即体积元 V_i 中的抽样点数 L_i 为

$$L_i = \frac{V_i}{V_s} L.$$

这里我们暂时将总抽样点数 L 作为一个固定值, 因为 L 的大小决定了求积分所需要的计算机时间.

在每一个体积元中利用期望值估计法求积分值及其方差 (式 (14.4.22)、式 (14.4.28))

$$I_i \approx V_i \hat{g}_i, \quad \sigma_i^2 \approx \frac{1}{L_i}\left[\hat{g}_i^2 - (\hat{g}_i)^2\right], \tag{14.4.42}$$

其中 $\hat{g}_i$ 是体积元 V_i 中被积函数的平均值

$$\left.\begin{aligned} \hat{g}_i &= \frac{1}{L_i}\sum_{j=1}^{L_i} g\left(\boldsymbol{\xi}_j\right) \\ \hat{g}_i^2 &= \frac{1}{L_i}\sum_{j=1}^{L_i} g^2\left(\boldsymbol{\xi}_j\right) \end{aligned}\right\} i = 1, 2, \cdots, N. \tag{14.4.43}$$

$\boldsymbol{\xi}_j (j = 1, 2, \cdots, L_i)$ 是体积元 V_i 中均匀分布的随机数. 于是积分值 I 及其误差可表示成

$$I = \sum_{i=1}^{N} I_i, \qquad \sigma^2 = \sum_{i=1}^{N} \sigma_i^2. \tag{14.4.44}$$

这就得到了积分及其方差的第一次迭代值. 为了将迭代进行下去, 可以采取两种不同的途径 (总的抽样点数 L 保持不变)

(1) 固定体积元 $V_i (i = 1, 2, \cdots, N)$ 不变, 调整每个体积元中的抽样点数 L_i.

(2) 固定原抽样点数 L_i 不变, 调整体积元 V_i 的大小.

下面我们仅叙述第一种途径的迭代步骤. 第二种途径比较复杂, 这里不予介绍, 有兴趣的读者可参阅文献 [110]、[111].

在第一次迭代计算 I_i 和 σ_i^2 的同时, 计算

$$\left.\begin{aligned} K_i \equiv \int_{V_i} |g(\boldsymbol{x})|\mathrm{d}\boldsymbol{x} \approx V_i|\hat{g}_i| \\ |\hat{g}_i| = \frac{1}{L_i}\sum_{j=1}^{L_i}|g(\boldsymbol{\xi}_j)| \end{aligned}\right\} i = 1, 2, \cdots, N \tag{14.4.45}$$

以及

$$K = \sum_{i=1}^{N} K_i,$$

在下次迭代中, 每个体积元中抽样点由原来的 L_i 改变为 L_i'

$$L_i' = \frac{K_i}{K}L. \tag{14.4.46}$$

这样, 每个体积元中的抽样点密度成为 (式 (14.4.41))

$$P_i' = \frac{1}{L}\frac{K_i}{K}\frac{L}{V_i} = \frac{|\hat{g}_i|}{K}, \qquad i = 1, 2, \cdots, N.$$

$|\hat{g}_i|$ 是 $|g(\boldsymbol{x})|$ 在体积元 V_i 中的平均值, 当 V_i 足够小, $|\hat{g}_i| \sim |g_i(\boldsymbol{x})|$, 因此, 近似地有

$$P_i' \approx \frac{|g_i(\boldsymbol{x})|}{\displaystyle\int_{V_s} |g(\boldsymbol{x})|\mathrm{d}\boldsymbol{x}}.$$

与式 (14.4.32) 对比可知, P_i' 接近于重要抽样方法中的概率密度 $f(\boldsymbol{x})$, 这就使积分估计值的方差大大减小. 如此迭代下去, 直至达到所要求的精度为止.

如果相当多次迭代之后尚不能达到所要求的精度, 则应考虑增加总的抽样点数目 L.

14.5 蒙特卡罗法应用于粒子传播问题

核物理、基本粒子物理研究的基本问题是原子核和基本粒子的衰变、相互作用, 这些过程中的一大部分都具有随机的性质, 例如, 不稳定核或粒子衰变产生一个或多个末态粒子, 这些末态粒子的种类、空间角分布、能量分布等等, 都是随机变量.

一个粒子射入介质时, 要经历一系列碰撞或相互作用, 其中的每一次碰撞或相互作用都是随机过程, 而且下一次碰撞与前一次碰撞是完全独立的. 这一连串的随机过程构成马尔科夫过程.

粒子穿入介质后的状况可以通过物理实验加以测定. 但如果粒子与物质的碰撞或相互作用的规律是清楚的, 那么粒子运动的这种马尔科夫过程, 完全能够用蒙特卡罗法正确地加以模拟. 通过对粒子的追踪, 可以从粒子的初态参数 (种类、能量、方向等) 得到任何中间态和末态的参数, 从而求得对于各种有关问题的答案.

这类涉及粒子的传播、粒子与物质的相互作用的问题可以统称为粒子传播问题, 它在基本粒子物理、核物理和宇宙线研究中是具有典型意义的.

粒子传播问题的经典处理方法是利用粒子传播的宏观方程来求得问题的解, 例如中子物理中的中子扩散方程. 宏观方程的优点在于对所研究的问题提供比较多的信息, 例如给出近似公式、渐近关系等, 因而对问题的解有整体的概貌性的了解. 但是对边界条件复杂的问题, 利用宏观方程很难处理; 更何况牵涉到多种过程的复杂问题, 常常无法求得宏观方程.

蒙特卡罗法恰恰相反, 只要问题所涉及的相互作用规律已经了解, 粒子的追踪是十分简单的, 任何复杂的边界条件都可以处理. 由于蒙特卡罗法对过程中的所有粒子都可以逐个追踪, 所以它可以给出非常详尽的信息, 例如, 任何时刻的状态, 任何界面上粒子的能量分布、飞行方向分布等, 这一优点是其他任何方法不能比拟的. 它的缺点在于只能给出描述问题的特定初态参数值下的具体答案, 而不能给出问题解的整体性质.

粒子传播问题中粒子的模拟追踪是根据粒子相互作用的物理规律的概率模型. 反过来, 如果对于某一种相互作用的物理规律不了解, 我们可以设想描写它的作用规律的模型, 利用蒙特卡罗法的计算结果与实验结果的对比来验证模型的正确性. 这种方法在粒子物理的研究中被广泛地采用.

蒙特卡罗法解粒子传播问题的要点可以归结为：①构造物理过程的概率分布并产生其随机数; ②处理边界条件; ③对于复杂的、尤其是涉及多分支过程的问题, 正确地组织和编写计算机程序.

蒙特卡罗法解粒子追踪问题显然涉及问题本身的物理内容. 为了避免过多地陷进物理本身的讨论, 我们以一个简化了的物理实验为例, 以说明蒙特卡罗法模拟物理实验的基本思路.

设想如图 14.12 的一个实验装置. 能量单一的准直 γ 光子细束垂直地射入被称为转换体的一块介质. γ 光子与物质的作用基本上是光电效应、康普顿散射和电子对产生, 我们只考虑康普顿散射, 即光子和原子中的一个电子碰撞, 作用结果是光子被散射偏离原来的运动方向、能量降低, 而电子得到一部分能量飞离原来所在的原子. 散射光子 $\gamma\prime$ 和电子 e 的飞出角度、能量有一定的分布. 如果图中的探测器

1(对 γ 光子灵敏) 测到散射光子 $\gamma\prime$, 探测器 2(对电子灵敏) 测到电子 e, 那么这两个探测器都有电脉冲输出, 符合电路就产生一个输出脉冲, 使得门电路打开, 探测器 2 的主脉冲 (其脉冲幅度正比于电子在探测器中的能量损失) 得以通过门电路并被脉冲幅度分析器记录下脉冲幅度. 当许多同样能量的光子射入转换体, 就能得到有一定分布的脉冲幅度谱, 称为响应谱. 利用一系列不同能量但数量相等的 γ 光子入射得到一系列的响应谱. 如果能量连续、强度未知的 γ 射线射入转换体, 得到对应的脉冲幅度谱, 则可以用测到的一系列响应谱对它进行分析, 求出入射 γ 射线的能量和强度分布.

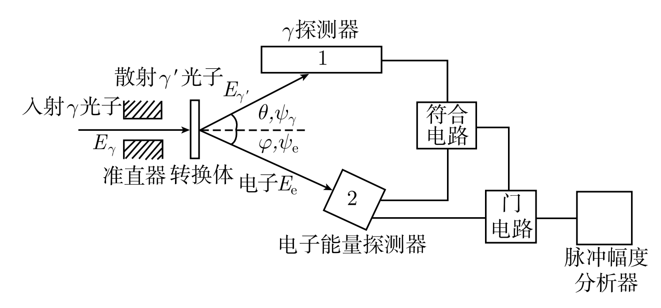


图 14.12　利用康普顿散射原理测量 γ 光子能谱的装置

响应谱自然可以利用实验来测定. 但是一系列不同能量的单能 γ 束是很难得到的. 我们可以通过蒙特卡罗法来模拟整个实验, 以求得任意单能 γ 束的响应谱.

当能量 E_γ 的光子垂直入射到转换体上, 在其中发生康普顿散射的概率为

$$p = 1 - \exp[-\sigma(E_\gamma)\cdot d], \qquad \sigma = \frac{z\rho N_a}{A}\cdot\sigma_a = C\sigma_a(E_\gamma), \tag{14.5.1}$$

其中 C 是常数, 与转换体物质种类有关; σ_a 是转换体物质原子的每个电子的康普顿散射截面, 与 E_γ 有关, 可由核数据表查到; d 是转换体厚度.

令康普顿散射后的散射光子和电子的能量分别为 E'_γ 和 E_e, 核物理知识告诉我们有下述关系：

$$E'_\gamma = \frac{E_\gamma}{1+\dfrac{E_\gamma}{0.511}(1-\cos\theta)}, \qquad E_\mathrm{e} = E_\gamma - E'_\gamma. \tag{14.5.2}$$

令散射光子对原入射光子的角度为 (θ,ψ_γ), 电子对原入射光子的角度为 $(\varphi,\psi_\mathrm{e})$, 其中 θ、φ 为极角, ψ_γ、ψ_e 为方位角, 则存在如下一一对应的关系 (图 14.12)：

$$\begin{aligned} &\mathrm{ctg}\varphi = \left(1+\frac{E_\gamma}{0.511}\right)\tan\frac{\theta}{2}, && 0\leqslant\theta,\varphi\leqslant\pi, \\ &\psi_\gamma = \psi_\mathrm{e}+\pi, && 0\leqslant\psi_\gamma,\psi_\mathrm{e}\leqslant 2\pi. \end{aligned} \tag{14.5.3}$$

如果其中的一个粒子 (如电子) 的飞行方向确定了, 则另一个的方向随之确定. 假定电子的飞出方向 $(\varphi, \psi_{\mathrm{e}})$ 的分布可用 $f(\varphi, \psi_{\mathrm{e}})$ 表示

$$\int_0^{2\pi}\int_0^{\pi} f(\varphi, \psi_{\mathrm{e}})\,\mathrm{d}\varphi\mathrm{d}\psi_{\mathrm{e}} = 1.$$

我们的问题是, 当 N_0 个能量 E_γ 的光子垂直射到转换体上, 脉冲幅度分析器接收到多少个脉冲, 它们的幅度分布 (谱) 如何.

当 N_0 个光子入射, 发生康普顿散射的光子数 N 由式 (14.5.1) 求出

$$N = N_0 p = N_0(1 - \mathrm{e}^{-\sigma d}).$$

这 N 次散射中, 每个散射电子的飞行方向由服从分布 $f(\varphi, \psi_{\mathrm{e}})$ 的 N 对随机数 $\{\xi_\varphi, \xi_{\psi_{\mathrm{e}}}\}$ 决定, 与之对应, 散射光子的飞行方向 θ^*, ψ_γ^* 由 $\xi_\varphi, \xi_{\psi_{\mathrm{e}}}$ 代入式 (14.5.3) 得出.

如果方向为 $\{\xi_\varphi, \xi_{\psi_{\mathrm{e}}}\}$ 的电子恰好落入探测器 2, 与之对应的散射光子又落入探测器 1(假定探测效率都是 100%), 则这是一个可被记录的有效事例. 若共有 n 个有效事例, 则脉冲幅度分析器接收到 n 个脉冲.

可以适当选择探测器 2 的形状和大小, 使得进入探测器 2 的电子被吸收全部能量, 这样探测器 2 的输出脉冲幅度 V 的分布为正态分布, 其均值 V_0 与电子能量成正比

$$g(V) \sim \exp\left\{-\frac{1}{2}\left(\frac{V - V_0}{\sigma_V}\right)^2\right\}, \qquad V_0 = bE_{\mathrm{e}}, \tag{14.5.4}$$

其中 b 是某个常数. 于是, n 个有效事例中的电子能量 E_{e} 由式 (14.5.2) 求得

$$E_{\mathrm{e}} = E_\gamma\left\{1 - \frac{1}{1 + \dfrac{E_\gamma}{0.511}(1 - \cos\theta^*)}\right\}.$$

这些电子在探测器 2 中产生的脉冲幅度由服从分布 $g(V)$ 的随机数 ξ_V 确定. 这 n 个 ξ_V 值就是我们要求的脉冲幅度谱, 即响应谱. 这样, 通过对实验过程中粒子的追踪, 得到了实验的 “模拟” 结果.

如果描述各个物理过程的公式是正确的, 在 N_0 很大的情形下, “模拟” 的结果将与实验结果相当一致. 由于模拟可以用计算机进行, 描述实验的几何和物理参数可以随意改变, 因此, 蒙特卡罗法模拟有很大的灵活性, 特别是由于条件的限制, 一些实验测量无法实行时, 蒙特卡罗法模拟 “实验” 的数据更有其特殊的重要性, 成为实验设计和实验结果分析的有力工具.

近代核物理、基本粒子物理的大型实验研究, 其设备庞大、精细、昂贵, 物理过程的相互作用复杂, 实验过程延续多年, 耗费大量人力和物力, 因此, 实验设计的优化极其重要, 而纯粹的解析方法几乎不能对实验的设计提供较为精确的预测. 同时, 对牵涉复杂过程的实验数据, 纯粹的解析方法同样不能给以恰当的说明. 因此, 蒙特卡罗法模拟成为核物理、基本粒子物理实验中进行实验设计和实验数据分析的最广泛使用、最重要的数学工具.

第 15 章　极小化方法

15.1　引　　言

在科学研究和实验测量的数据处理中, 经常遇到最优化问题, 它所研究的问题是在众多可供选择的方案中选择最优方案. 最优方案通常是使描述物理问题的一个特定函数 —— 称为目标函数 —— 达到极大或极小. 例如, 我们前面讨论了利用测量数据对理论模型的未知参数进行估计的极大似然法和最小二乘法, 就是认为使似然函数达到极大和使最小二乘函数 Q^2 达到极小的参数值是最优估计值. 函数的求极小与极大是完全相当的运算, 对目标函数 $f(\boldsymbol{x})$ 求极大, 相当于对目标函数 $-f(\boldsymbol{x})$ 求极小. 最优化问题有些书籍称为极值问题或数学规划问题, 本书中采用极小化这一称呼.

最优化方法已发展成为应用数学的一个专门分支, 希望深入了解这一方法的读者可参阅有关的专门书籍和文献. 鉴于极小化方法在极大似然法和最小二乘法做参数估计中的关键性作用, 在本章中我们选择较有代表性的算法做一简略的介绍.

函数极小化问题一般地可表述为求**目标函数**$f(\boldsymbol{x})$ 达到极小时自变量 $\boldsymbol{x}$ 的值 $\boldsymbol{x}^*$, 即

$$\min f(\boldsymbol{x}) = f(\boldsymbol{x}^*), \qquad \boldsymbol{x} \in E^n, \tag{15.1.1}$$

并满足约束条件方程

$$h_i(\boldsymbol{x}) = 0, \qquad i = 1, 2, \cdots, l, \qquad l < n, \tag{15.1.2}$$

$$g_j(\boldsymbol{x}) \geqslant 0, \qquad j = 1, 2, \cdots, m, \tag{15.1.3}$$

其中 $\boldsymbol{x} \in E^n$ 表示 $\boldsymbol{x}$ 是 n 维欧氏 (即欧几里得) 空间中的一个点, 或者说是 n 维向量, $x_1, x_2, \cdots, x_n$ 是 $\boldsymbol{x}$ 的 n 个分量. $f(\boldsymbol{x})$, $h_i(\boldsymbol{x})$, $g_j(\boldsymbol{x})$ 都是 $\boldsymbol{x}$ 的函数. 如果不存在约束条件, 即 $l, m = 0$, 称为无约束极小化问题; 反之则称为约束极小化问题. 式 (15.1.2) 称为等式约束; 式 (15.1.3) 称为不等式约束. 对于 $g'_j(\boldsymbol{x}) \leqslant 0$ 这类不等式约束, 可令 $g_j(\boldsymbol{x}) = -g'_j(\boldsymbol{x})$ 化为式 (15.1.3) 的形式, 因此, 式 (15.1.2) 和式 (15.1.3) 的约束方程具有一般性.

在极大似然法中, 若 $\boldsymbol{\vartheta}$ 是待估计的未知参数, 则其极大似然估计值 $\hat{\boldsymbol{\vartheta}}$ 满足

$$\min\,[-L(\boldsymbol{\vartheta})] = -L(\hat{\boldsymbol{\vartheta}});$$

而在最小二乘法中, 则应有

$$\min Q^2(\boldsymbol{\vartheta}) = Q^2(\hat{\boldsymbol{\vartheta}}).$$

与式 (15.1.1) 对照知道, 本章中的 $\boldsymbol{x}$ 相当于参数估计问题中的待估计未知参数, 而 $\boldsymbol{x}^*$ 则是未知参数的最优估计值. 至于未知参数服从某种约束方程的问题, 在 9.7 节, 9.8 节中已作过简略的讨论.

从数学分析知道, 若目标函数 $f(\boldsymbol{x})$ 存在连续的一阶偏导数, 变量值 $\boldsymbol{x}^*$ 为 $f(\boldsymbol{x})$ 稳定点的必要条件是函数在 $\boldsymbol{x}^*$ 点的梯度向量等于 $\mathbf{0}$, 即

$$\nabla f(\boldsymbol{x}^*) = \left(\frac{\partial f(\boldsymbol{x})}{\partial x_1}, \frac{\partial f(\boldsymbol{x})}{\partial x_2}, \cdots, \frac{\partial f(\boldsymbol{x})}{\partial x_n}\right)^{\mathrm{T}}_{x=x^*} = \mathbf{0}. \tag{15.1.4}$$

为了对目标函数极小化, 它在某一点 $\boldsymbol{x}_0$ 附近的行为利用泰勒级数来逼近

$$\begin{aligned} f(\boldsymbol{x}) \approx & f(\boldsymbol{x}_0) + (\boldsymbol{x} - \boldsymbol{x}_0)^{\mathrm{T}} \nabla f(\boldsymbol{x}_0) \\ & + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0)^{\mathrm{T}} \nabla^2 f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + \cdots, \end{aligned} \tag{15.1.5}$$

其中 $\nabla^2 f(\boldsymbol{x}_0)$ 是目标函数 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}_0$ 点的二阶导数矩阵, 也称为**黑塞 (Hesse) 矩阵**. 这是一个 $n \times n$ 对称方矩阵, 其元素是

$$\left.\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j}\right|_{x=x^*}, \qquad i, j = 1, 2, \cdots, n. \tag{15.1.6}$$

式 (15.1.5) 右边级数的第一项是常数, 对极小点的确定不能提供任何信息. 第二项是函数在 $\boldsymbol{x}_0$ 点的梯度, 它指示出函数在该点附近上升或下降最快的方向; 当函数处于稳定点时, 其梯度值为 0, 即如式 (15.1.4) 所示. $\boldsymbol{x}^*$ 成为 $f(\boldsymbol{x})$ 的极小点的充分条件是除了式 (15.1.4) 成立之外, 还要求 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}^*$ 点的黑塞矩阵为**正定矩阵**. 所谓矩阵 $\underset{\sim}{\boldsymbol{A}}$ 为正定, 是指对于任何向量 $\boldsymbol{x} \neq \mathbf{0}$, 总有

$$\boldsymbol{x}^{\mathrm{T}} \underset{\sim}{\boldsymbol{A}} \boldsymbol{x} > 0 \tag{15.1.7}$$

通过黑塞矩阵的正定性来判断稳定点是极小点在数学上是很重要的, 但在许多实际问题中, 由于种种原因 (如目标函数无法用解析函数表示, 二阶偏导数难以求出等) 很难实现, 这时可从所得的解就问题本身作出判断.

由以上介绍可知, 求函数极小值的问题可化为求解

$$\nabla f(\boldsymbol{x}) = \mathbf{0} \tag{15.1.8}$$

的问题, 这是含 n 个未知变量、n 个方程组成的方程组. 除了目标函数是二次函数、该方程组是线性方程组, 可以求得解析解之外, 一般情形下这是非线性方程组, 很难用解析方法求解. 有时, 目标函数的行为可能十分复杂和奇特, 甚至无法写出一

阶导数的解析形式. 所以求极小值一般用数值方法, 其中最常见的是迭代法, 其基本思想如下：首先给出极小值点的估计初值 $\boldsymbol{x}_0$, 然后根据某种算法计算一系列的 $\boldsymbol{x}_k(k=1,2,\cdots)$, 使得点列 $\{x_k\}$ 的极限就是 $f(\boldsymbol{x})$ 的一个极小值点 $\boldsymbol{x}^*$. 在实际运算时需要给定一个判据, 当迭代进行到满足该判据时终止迭代, 求出的变量值认为是极小值点 $\boldsymbol{x}^*$ 的近似值. 终止迭代的判据要视具体的算法和问题的性质而定, 常用的判据可以是两次相连迭代中函数值之差小于给定的正数

$$\|f(\boldsymbol{x}_{k+1})-f(\boldsymbol{x}_k)\|\leqslant\varepsilon,$$

或者梯度的长度小于给定的正数

$$\|\nabla f(\boldsymbol{x}_k)\|\leqslant\varepsilon,$$

其中符号 $\|z\|$ 称为变量 z 的**模**或**长度**, 其定义是

$$\|z\|=\sqrt{\sum_{j=1}^{n}z_j^2}. \tag{15.1.9}$$

求得目标函数 $f(\boldsymbol{x})$ 的近似极小值点 $\boldsymbol{x}_k^*$ 后, 余下的问题是确定 $\boldsymbol{x}_k^*$ 对于实际极小值点 $\boldsymbol{x}^*$ 的误差, 它可由 $\boldsymbol{x}$ 的协方差矩阵及求出的 $\boldsymbol{x}_k^*$ 求得. 如果所采用的极小化方法的迭代步骤中求出了黑塞矩阵在近似极小点值 $\boldsymbol{x}_k^*$ 处的值, 则 $\boldsymbol{x}$ 的协方差矩阵可由 $\underset{\sim}{H}^{-1}(\boldsymbol{x}_k^*)$ 表示 (见 15.8 节), 这里 $\underset{\sim}{H}(\boldsymbol{x})$ 表示 $\boldsymbol{x}$ 的黑塞矩阵. 如果极小化方法迭代步骤中没有做黑塞矩阵的计算, 则必须对目标函数做一些合理的假定来进行协方差矩阵的计算.

由以上所述可知, 极小化方法的核心在于找到一种计算 $\boldsymbol{x}_k$ 的算法, 使得点列 $\{\boldsymbol{x}_k\}$ 能逼近目标函数 $f(\boldsymbol{x})$ 的极小点 $\boldsymbol{x}^*$, 逼近的速度越快, 算法的效率就越高. 迭代过程需要多次重复计算不同点 $\boldsymbol{x}_k$ 处的目标函数值, 甚至导数值, 这种大量重复的计算不利用高速电子计算机事实上是无法完成的. 各种极小化方法的效率, 可以用对于同一个问题求出达到同样精确度的解所需的计算机机时来衡量. 需要注意的是, 计算机由于字长的限制导致数字的舍入误差, 由于计算机对整数和实数的表示存在上界和下界, 因此存在 “上溢”(大于计算机最大可表示的数判为无穷大) 和 “下溢”(小于计算机最小可表示的数判为零) 的现象. 在极小法算法中必须考虑到它们对算法的精度、效率、极小值点解的收敛性、稳定性等造成的影响.

15.2　无约束极小化的一维搜索

我们首先介绍无约束极小化的一维搜索方法, 这是后面各节 n 维极小化问题求解的基础. 一维搜索问题就是求解方程

$$\min f(x)=f(x^*),\qquad -\infty<x<+\infty,$$

或写为

$$\min f(x) = f(x^*), \qquad x \in E^1. \tag{15.2.1}$$

由极值的必要条件可知, 极小值点必为稳定点, 故一维极小化问题又可表为求解方程

$$f'(x) = 0. \tag{15.2.2}$$

各种一维搜索方法通过求解式 (15.2.1) 或式 (15.2.2), 得到近似极小值点. 一维搜索方法可以分为两类, 一类需要计算目标函数的导数, 它们一般收敛向极小值点的速度较快, 但对不易求导的复杂函数及设有解析表达式的函数无法使用; 另一类只需要计算点列 $\{x_R\}$ 的目标函数值, 因而通用性较强, 但一般说来收敛向极小值点的速度不如前一类快, 这类方法可称为直接搜索. 本节将介绍的黄金分割法、斐波那契法、二次函数插值法和成功–失败法等四种方法属于直接搜索法; 属于第一类的牛顿法在 n 维无约束极小化方法中介绍, 因为对于 $n = 1$ 的特殊情形同样适用.

15.2.1 黄金分割法 (0.618 法)

假定目标函数在所考虑的区间 $[a,b]$ 内是单峰函数, 即在 $[a,b]$ 内只存在一个局部极小点. 在 $[a,b]$ 内任选两点 $x_1, x_1'(x_1 < x_1')$. 若 $f(x_1) < f(x_1')$; 则极小值点 x^* 必在区间 $[a, x_1']$ 内; 反之, x^* 必在 $[x_1, b]$ 内. 所以比较 $f(x_1)$ 和 $f(x_1')$ 的大小可以确定 x^* 所在的区间, 将该区间记为 $[a_1, b_1]$. 重复上述步骤可得到新的区间 $[a_2, b_2]$, $[a_3, b_3], \cdots$, 使包含 x^* 的区间不断缩小. 当迭代到第 k 次, 满足 $|b_k - a_k| < \varepsilon, \varepsilon$ 为给定正数时, 可停止迭代, 取

$$x_k^* = \frac{a_k + b_k}{2}$$

作为 x^* 的近似最优解. 由于区间 $[a_k, b_k]$ 包含极小值点 x^*, 显然有

$$|x_k^* - x^*| \leqslant \varepsilon/2,$$

即解 x_k^* 的绝对误差 $\varepsilon_0 \leqslant \varepsilon/2$.

问题是如何选取点列 $\{x_i, x_i'\}$ 才能减少运算量. 每次选定了 x_i 和 x_i', 并比较 $f(x_i)$ 和 $f(x_i')$ 后可求得包含 x^* 的长度减小的新区间, 如果在 x_i 和 x_i' 中的一个点在下一次迭代中仍能使用, 则每步只需计算一次目标函数值, 从而节省了一半运算. 达到这一目的的方法如下：

如图 15.1 所示, 在长度 t_0 的单峰区间 $[a,b]$ 内选定两点 x_1, x_1', 使得线段 ax_1' 和 x_1b 长度都等于 $t_1, (t_1 > t_0/2)$ 并且 $x_1 < x_1'$. 这样, 不论 $f(x_1') > f(x_1)$ 还是 $f(x_1') < f(x_1)$, 包含 x^* 的区间 ax_1' 或 x_1b 长度都是 t_1, 即包含 x^* 区间的长度由 t_0 缩小到 t_1, 或者说经过一次迭代运算的**区间缩短因子**$\alpha = t_1/t_0$. 不失一般性, 假定 $f(x_1') > f(x_1)$, 则 x^* 落在区间 $[a,x_1']$ 内. 在该区间内选择两点 x_2, x_2', 满足

$x_2 < x_2', \overline{ax_2'} = \overline{x_2x_1'} = t_2(t_2 > t_1/2)$. 为了使第一次迭代中选出的 x_1 在本次迭代中能继续使用, 令 $x_2' = x_1$, 由图 15.1 可知, $t_2 = t_0 - t_1$. 为了使区间缩短因子在本次迭代与上次迭代中相同, 应有 $\alpha = t_2/t_1 = t_1/t_0$. 由这些关系可求出

$$\begin{cases} \alpha = \left(\sqrt{5}-1\right)/2 \approx 0.618, \\ x_i = b_i - \alpha(b_i - a_i), \\ x_i' = a_i + \alpha(b_i - a_i). \end{cases} \tag{15.2.3}$$

其中 a_i, b_i 表示第 i 次迭代中包含极小值点 x^* 的区间, x_i 和 x_i' 表示该次迭代应选的两点的坐标值, α 是一次迭代的区间缩短因子, 由于它的值近似地等于 0.618, 所以也称为 0.618 方法. 式 (15.2.3) 唯一地确定了黄金分割法中迭代点的选取. 若共迭代了 $n-1$ 次, 共计算目标函数值 n 次 (第一次迭代需计算两次目标函数值), 最后得到包含 x^* 的区间长度为

$$|b_{n-1} - a_{n-1}| = (b-a)0.618^{n-1} \leqslant \varepsilon,$$

定义黄金分割法的效率 (经过 n 次函数值计算) 为

$$E_n = \frac{b-a}{b_{n-1} - a_{n-1}} = 0.618^{-(n-1)}, \tag{15.2.4}$$

显然有

$$n-1 \geqslant \frac{\lg\left(\dfrac{\varepsilon}{b-a}\right)}{\lg 0.618} \geqslant \frac{\lg\left(\dfrac{2\varepsilon_0}{b-a}\right)}{\lg 0.618}. \tag{15.2.5}$$

当 x^* 的近似解的误差要求 ε_0 确定之后, 由式 (15.2.5) 可求得必要的迭代次数 $n-1$.

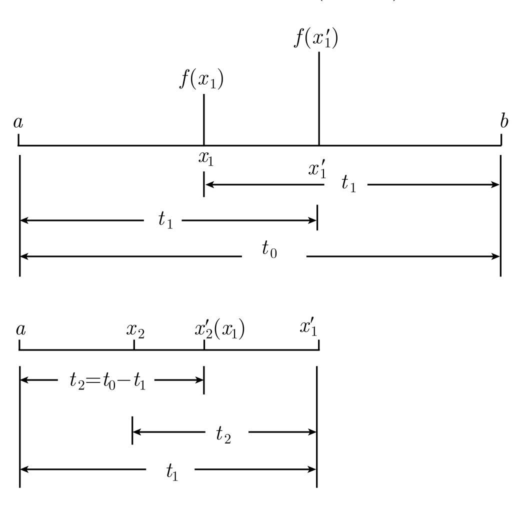


图 15.1　一维搜索的黄金分割法

图 15.2 给出黄金分割法计算近似极小值点的程序框图, 由此容易编制计算机程序.

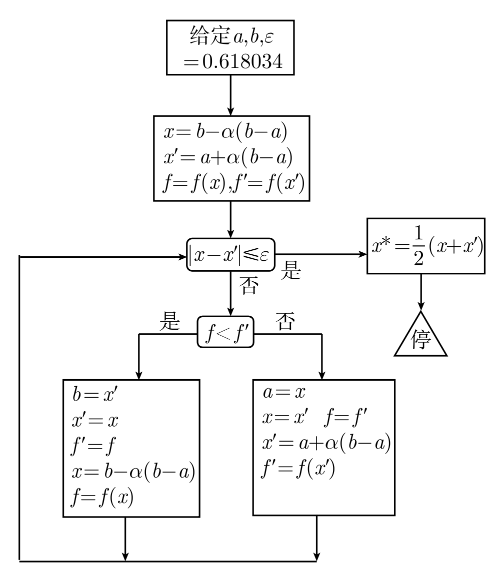


图 15.2 黄金分割法程序框图

15.2.2 斐波那契法

在某些实际问题中, 要求限定迭代的次数. 显然, 对于一定的迭代次数, 如果某种极小化方法确定的近似极小值点误差最小, 这种方法是比较理想的. 斐波那契法符合这一要求.

由 15.2.1 节所述可知, 0.618 法有两个特点. 一个是每次迭代只需计算一次目标函数值, 使计算量减少了一半, 这一点应当保留; 其次是每迭代一次包含极小值点的区间缩短因子等于常数 0.618, 这一点未必见得符合最佳方案.

在斐波那契搜索中, 用到所谓斐波那契数列 F_k, 其定义是

$$F_0=F_1=1, \qquad F_k=F_{k-1}+F_{k-2}, \qquad k=2,3,\cdots. \tag{15.2.6}$$

设 $f(x)$ 为区间 $[a,b]$ 上的单峰函数, 规定共进行 n 次函数值计算. 在斐波那契法中, 第一次迭代取

$$x_1=\frac{F_{n-2}}{F_n}(b-a)+a, \qquad x_1'=\frac{F_{n-1}}{F_n}(b-a)+a.$$

比较 $f(x_1)$ 和 $f(x_1')$ 值的大小, 即可知道包含极小值点的区间是 $[a,x_1']$ 还是 $[x_1,b]$. 设该区间为 $[a,x_1']$, 此时, 区间缩短因子为

$$\alpha_1 = \frac{[a, x_1']}{[a, b]} = \frac{x_1' - a}{b - a} = \frac{F_{n-1}(b-a)}{F_n(b-a)} = \frac{F_{n-1}}{F_n},$$

或包含极小值点区间是 $[x_1, b]$, 则

$$\begin{aligned}\alpha_1 = \frac{[x_1, b]}{[a, b]} = \frac{b - x_1}{b - a} &= \frac{b - \dfrac{F_{n-2}}{F_n}(b-a) - a}{b - a} \\ &= 1 - \frac{F_{n-2}}{F_n} = \frac{F_{n-1}}{F_n}.\end{aligned}$$

记缩短了的区间为 $[a_1, b_1]$, 依照上述方法取两个插入点 x_2 和 x_2' 为

$$x_2 = \frac{F_{n-3}}{F_{n-1}}(b_1 - a_1) + a_1, \qquad x_2' = \frac{F_{n-2}}{F_{n-1}}(b_1 - a_1) + a_1.$$

为了使每次迭代只计算一个函数值, 要求下列情况之一发生：

$$x_1 = x_2, \qquad x_1' = x_2, \qquad x_1 = x_2', \qquad x_1' = x_2'.$$

我们来验证这一点是否成立. 当 $f(x_1) \leqslant f(x_1')$ 时, $a_1 = a, b_1 = x_1'$, 于是

$$x_2' = \frac{F_{n-2}}{F_{n-1}}(b_1 - a_1) + a_1 = \frac{F_{n-2}}{F_{n-1}}(x_1' - a) + a = \frac{F_{n-2}}{F_n}(b-a) + a = x_1,$$

即 x_2' 与 x_1 是同一个点. 而当 $f(x_1') < f(x_1)$ 时, $a_1 = x_1, b_1 = b$, 故

$$\begin{aligned}x_2 &= \frac{F_{n-3}}{F_{n-1}}(b_1 - a_1) + a_1 = \frac{F_{n-3}}{F_{n-1}}(b - x_1) + x_1 \\ &= \frac{F_{n-3}}{F_{n-1}}\left[b - \frac{F_{n-2}}{F_n}(b-a) - a\right] + \frac{F_{n-2}}{F_n}(b-a) + a \\ &= \frac{F_{n-3}}{F_{n-1}}(b-a)\left[1 - \frac{F_{n-2}}{F_n}\right] + \frac{F_{n-2}}{F_n}(b-a) + a \\ &= \frac{F_{n-3}}{F_{n-1}}(b-a)\frac{F_{n-1}}{F_n} + \frac{F_{n-2}}{F_n}(b-a) + a \\ &= (b-a)\frac{F_{n-2} + F_{n-3}}{F_n} + a = (b-a)\frac{F_{n-1}}{F_n} + a = x_1',\end{aligned}$$

即 x_2 与 x_1' 是同一个点. 区间缩短因子为

$$\alpha_2 = \frac{[a_1, x_2']}{[a_1, b_1]} = \frac{x_2' - a_1}{b_1 - a_1} = \frac{\dfrac{F_{n-2}}{F_{n-1}}(b_1 - a_1)}{(b_1 - a_1)} = \frac{F_{n-2}}{F_{n-1}},$$

或

$$\alpha_2 = \frac{[x_2, b_1]}{[a_1, b_1]} = \frac{b_1 - x_2}{b_1 - a_1} = \frac{b_1 - \dfrac{F_{n-3}}{F_{n-1}}(b_1 - a_1) - a_1}{b_1 - a_1}$$
$$= 1 - \frac{F_{n-3}}{F_{n-1}} = \frac{F_{n-2}}{F_{n-1}}.$$

利用数学归纳法可以证明, 在第 $i+1$ 次迭代中, 第 i 次迭代中的一个点可以作为保留点

$$x'_{i+1} = x'_i = \frac{F_{n-i-1}}{F_{n-i}}(b_i - a_i) + a_i, \qquad \text{若} f(x_i) \leqslant f(x'_i) \tag{15.2.7}$$

或

$$x_{i+1} = x'_i = \frac{F_{n-i-2}}{F_{n-i}}(b_i - a_i) + a_i, \qquad \text{若} f(x_i) > f(x'_i). \tag{15.2.8}$$

因此, 每次迭代只需计算一个新的目标函数值. 第 i 次迭代中的区间缩短因子为

$$\alpha_i = \frac{F_{n-i}}{F_{n-i+1}}. \tag{15.2.9}$$

当进行到 $n-2$ 次迭代时, 根据式 (15.2.7)、式 (15.2.8), 有

$$x_{n-2} = a_{n-3} + \frac{F_1}{F_3}(b_{n-3} - a_{n-3}) = a_{n-3} + \frac{1}{3}(b_{n-3} - a_{n-3}),$$

$$x'_{n-2} = a_{n-3} + \frac{F_2}{F_3}(b_{n-3} - a_{n-3}) = a_{n-3} + \frac{2}{3}(b_{n-3} - a_{n-3}).$$

计算 x_{n-2} 和 x'_{n-2} 点的目标函数值并作比较, 分两种情况:

(1) $f(x_{n-2}) > f(x'_{n-2})$, 包含极小值点的区间是 $[a_{n-2}, b_{n-2}] = [x_{n-2}, b_{n-3}]$, 保留点是

$$x'_{n-2} = a_{n-3} + \frac{2}{3}(b_{n-3} - a_{n-3}) = \frac{a_{n-2} + b_{n-2}}{2};$$

(2) $f(x_{n-2}) \leqslant f(x'_{n-2})$, 含极小值点区间为 $[a_{n-2}, b_{n-2}] = \left[a_{n-3}, x'_{n-2}\right]$, 保留点是

$$x_{n-2} = a_{n-3} + \frac{1}{3}(b_{n-3} - a_{n-3}) = \frac{a_{n-2} + b_{n-2}}{2}.$$

可见, 第 $n-2$ 次迭代完成时, 保留点都等于区间 $[a_{n-2}, b_{n-2}]$ 的中点, 记为 $x_{n-2}^{(m)}$. 这时, 总共已做了 $n-1$ 次目标函数值的计算 (第一次迭代计算两个函数值, 其余迭代各计算一次). 在第 $n-1$ 次迭代中, 选取

$$x_{n-1} = x_{n-2}^{(m)} + \delta, \qquad 0 < \delta < \frac{b_{n-2} - a_{n-2}}{2}.$$

计算 $f(x_{n-1})$ 并与 $f(x_{n-2}^{(m)})$ 比较, 得到近似极小值点 x_n^* 为

$$x_n^* = \begin{cases} \dfrac{1}{2}\left(x_{n-2}^{(m)} + b_{n-2}\right), & f(x_{n-1}) < f(x_{n-2}^{(m)}), \\ \dfrac{1}{2}\left(x_{n-2}^{(m)} + a_{n-2}\right), & f(x_{n-1}) \geqslant f(x_{n-2}^{(m)}), \end{cases} \tag{15.2.10}$$

这样确定的 x_n^* 与真实极小值点 x^* 的误差 ε_0 为

$$\varepsilon_0 \equiv |x_n^* - x^*| \leqslant \frac{1}{2}\left(b_{n-2} - a_{n-2}\right).$$

由式 (15.2.9) 可求得

$$\begin{aligned}(b_{n-2} - a_{n-2}) &= (b_{n-3} - a_{n-3})\frac{F_2}{F_3} = (b_{n-4} - a_{n-4})\frac{F_3}{F_4}\frac{F_2}{F_3} = \cdots \\ &= (b-a)\frac{F_{n-1}}{F_n}\cdot\frac{F_{n-2}}{F_{n-1}}\cdot \ \cdots \ \cdot\frac{F_3}{F_4}\cdot\frac{F_2}{F_3} = (b-a)\frac{F_2}{F_n} \\ &= 2(b-a)/F_n.\end{aligned}$$

因此

$$\varepsilon_0 \equiv |x_n^* - x^*| \leqslant \frac{b-a}{F_n}. \tag{15.2.11}$$

在搜索区间 $[a,b]$ 和目标函数的计算次数 n 确定之后, 式 (15.2.11) 唯一地确定了最后求出的近似极小值点 x_n^* 的误差; 反过来, 也可以由要求的精度 ε_0 求出所需的迭代次数 $n-1$.

经过 $n-1$ 次迭代 (n 次目标函数值计算), 斐波那契法的效率为

$$E_n = \frac{b-a}{\dfrac{1}{2}(b_{n-2} - a_{n-2})} = F_n. \tag{15.2.12}$$

与黄金分割法的效率 $E_n = 0.618^{-(n-1)}$ 相比, 通过直接计算可知

$$F_n > (0.618)^{-(n-1)},$$

所以斐波那契搜索的效率高于 0.618 法.

利用数学归纳法可以证明：

$$F_k = \frac{1}{\sqrt{5}}\left\{\left(\frac{1+\sqrt{5}}{2}\right)^{k+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k+1}\right\},$$

故

$$\lim_{k\to\infty}\frac{F_{k-1}}{F_k} = \lim_{k\to\infty}\frac{\left(\dfrac{1+\sqrt{5}}{2}\right)^{k} - \left(\dfrac{1-\sqrt{5}}{2}\right)^{k}}{\left(\dfrac{1+\sqrt{5}}{2}\right)^{k+1} - \left(\dfrac{1-\sqrt{5}}{2}\right)^{k+1}}$$

$$= \frac{1}{\frac{1+\sqrt{5}}{2}} = \frac{\sqrt{5}-1}{2},$$

这正是黄金分割法中的区间缩短因子, 所以黄金分割法是斐波那契搜索的极限情形. 表 15.1 给出 F_{k-1}/F_k 数列.

表 15.1 斐波那契数列 F_k 和区间缩短因子 F_{k-1}/F_k

k	F_k	F_{k-1}/F_k	k	F_k	F_{k-1}/F_k
1	1	1	9	55	0.618 18
2	2	0.5	10	89	0.617 98
3	3	0.666 67	11	144	0.618 06
4	5	0.6	12	233	0.618 026
5	8	0.625	13	377	0.618 037
6	13	0.615 38	14	610	0.618 033
7	21	0.619 05	⋮		⋮
8	34	0.617 65	∞		0.618 034

由表 15.1 可见, 当 $k>7$, 斐波那契法的区间缩短因子 F_{k-1}/F_k 的数值取小数点后三位都是 0.618. 因此, 尽管从理论上讲, 斐波那契法较为有效, 但因黄金分割法的计算程序简单得多, 而两者的效率相差不大, 故一般宁愿采用黄金分割法.

图 15.3 给出了斐波那契法的程序框图, 其中用到了下述关系式：

$$\begin{aligned} x_i + x_i' &= \frac{F_{n-i-1}}{F_{n-i+1}}(b_{i-1}-a_{i-1}) + a_{i-1} + \frac{F_{n-i}}{F_{n-i+1}}(b_{i-1}-a_{i-1}) + a_{i-1} \\ &= \frac{F_{n-i-1}+F_{n-i}}{F_{n-i+1}}(b_{i-1}-a_{i-1}) + 2a_{i-1} = b_{i-1} - a_{i-1} + 2a_{i-1} \\ &= b_{i-1} + a_{i-1}. \end{aligned} \tag{15.2.13}$$

15.2.3 二次函数插值法 (抛物线法)

二次函数插值法 (也称抛物线法) 是利用二次函数 $\varphi(x)$ 逼近目标函数 $f(x)$, 并取 $\varphi(x)$ 的极小值点作为 $f(x)$ 极小值点的近似, 从而构造出迭代算法, 直到得出符合终止迭代判据的结果为止. 本方法中, 二次函数是利用三个不同点上的目标函数值来构造的.

设已有三个点 x_1, x_2, x_3, 对应的目标函数值为 f_1, f_2, f_3. 通过三个点 (x_1, f_1), $(x_2, f_2), (x_3, f_3)$ 的抛物线 (二次函数) 方程是

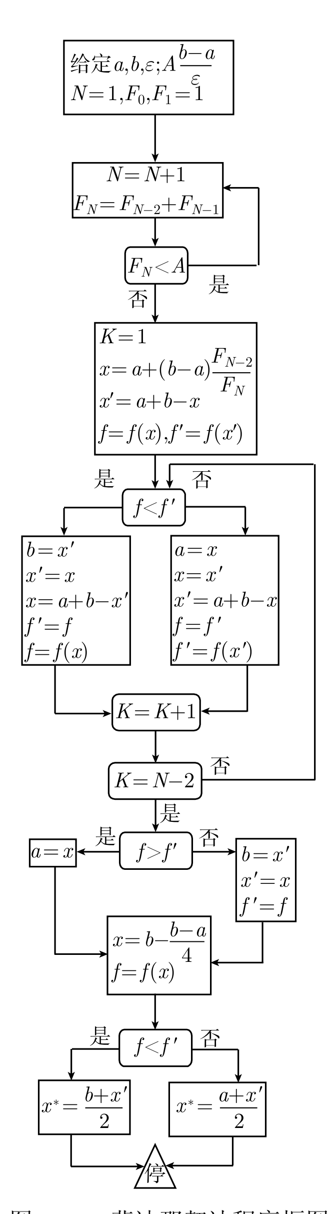


图 15.3　斐波那契法程序框图

$$\varphi(x)=\frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)}f_1+\frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)}f_2+\frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)}f_3.$$

令 $\varphi'(x)=0$, 得 $\varphi(x)$ 的极值点为

$$\overline{x}=\frac{1}{2}\frac{f_1\left(x_2^2-x_3^2\right)+f_2\left(x_3^2-x_1^2\right)+f_3\left(x_1^2-x_2^2\right)}{f_1(x_2-x_3)+f_2(x_3-x_1)+f_3(x_1-x_2)}.\tag{15.2.14}$$

当满足

$$x_1 < x_2 < x_3, \qquad f_1 > f_2 > f_3 \tag{15.2.15}$$

的条件时, $\overline{x}$ 为 $\varphi(x)$ 的极小值点. 式 (15.2.15) 相当于函数值两边高中间低. 特别当 x_1, x_2, x_3 为等间距的情形, $\overline{x}$ 的表达式简化为

$$\overline{x} = x_2 + \frac{d}{2}\,\frac{f_1 - f_3}{f_1 + f_3 - 2f_2}, \qquad x_2 - x_1 = x_3 - x_2 = d. \tag{15.2.16}$$

如果区间长度 $|x_3 - x_1|$ 足够小, 即对于给定常数 ε 有 $|x_3 - x_1| < \varepsilon$, 则 $\overline{x}$ 可作为 $f(x)$ 极小值点 x^* 的近似值, 两者的差别为

$$|x^* - \overline{x}| < |x_3 - x_1| < \varepsilon, \tag{15.2.17}$$

这时可终止迭代. 也可以根据目标函数值来决定迭代的终止

$$|f(\overline{x}) - f_2| < \varepsilon'. \tag{15.2.18}$$

当 $f(\overline{x}) < f_2$, 取 $\overline{x}$ 为 x^* 的近似值; 否则以 x_2 作为 x^* 的近似值.

如果迭代终止, 判据式 (15.2.17) 或 (15.2.18) 未得到满足, 则应寻找三个新的点 x_1', x_2', x_3' 继续做上述的计算, 直到求出满意的解为止. 为了能减少目标函数值的计算, x_1', x_2', x_3' 的选择应满足如下条件：

(1) $x_1' < x_2' < x_3', f_1' < f_2' < f_3'$;

(2) $x_3' - x_1' < x_3 - x_1$;

(3) x_1', x_2', x_3' 的选择应充分利用原有的四个点 $x_1, x_2, x_3, \overline{x}$ 的信息.

x_1', x_2', x_3' 的一种可能的方案表示在图 15.4 的二次插值法程序框图中.

对于一般的目标函数, 抛物线法一般地说比斐波那契法的收敛速度快; 特别当目标函数接近于二次函数时, 抛物线法能迅速地逼近极小值点. 它的不足之处是当三个点处于同一直线上, 即满足

$$f_1(x_2 - x_3) + f_2(x_3 - x_1) + f_3(x_1 - x_2) = 0 \tag{15.2.19}$$

时导致发散; 当三个点接近于同一直线上, 求出的极小值点可能偏离真正的极值点相当远. 但是一般地说, 在极小值点附近用抛物线来逼近目标函数是相当好的近似, 只要利用目标函数值满足式 (15.2.15) 的要求选择 x_1, x_2, x_3, 利用抛物线法求出近似极小值点是比较迅速和可靠的.

15.2.4 进退法

进退法又称成功–失败法, 它的基本思想如下：任选一个初始点 x_0 和初始步长 d_0, 取 $x_0 + d_0$ 为新的试探点, 计算目标函数值 $f_0 = f(x_0)$ 和 $f_1 = f(x_1) = f(x_0 + d_0)$.

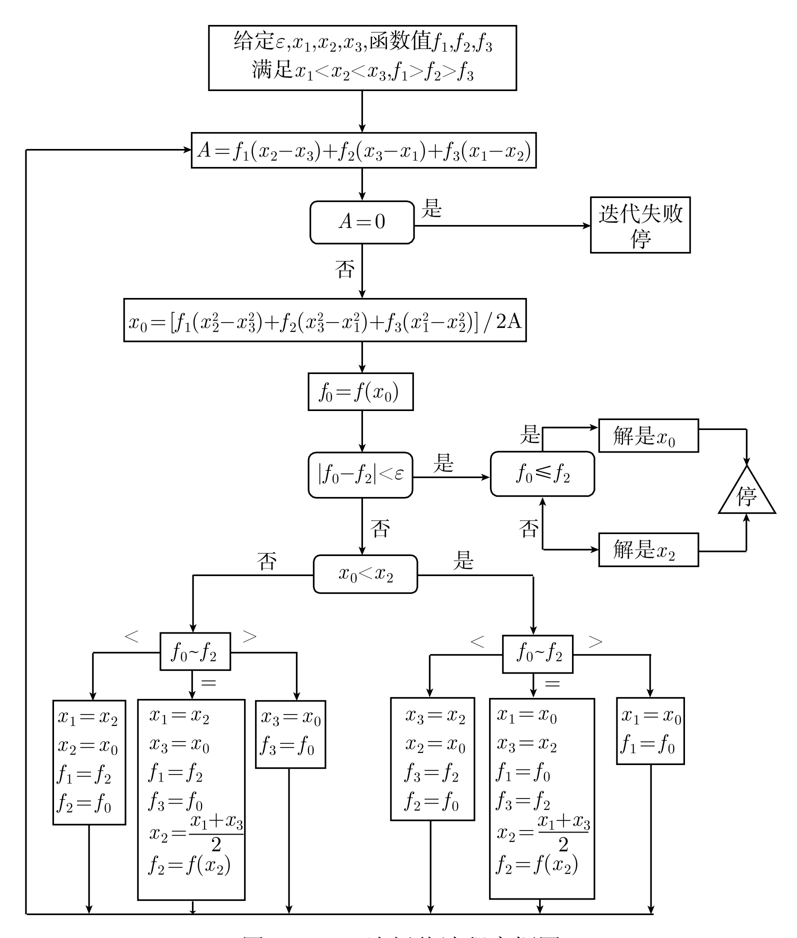


图 15.4 二次插值法程序框图

如果 $f_1 < f_0$, 则从 x_0 到 x_1 目标函数值下降, 称为试验成功; 反之, 则称为试验失败. 在试验成功的情形下, 搜索方向是函数下降方向, 可沿该方向增大步长继续搜索, 故可以 x_1 为新的初始点, $\alpha d_0(\alpha > 1)$ 作为新的步长 d_1 进行新的试验. α 称为扩展因子, 一般取 $\alpha \approx 3.0$. 如果试验失败, 则原搜索方向是目标函数增大的方向, 故应沿反向搜索, 以 x_0 为初始点, 以 $-\beta d_0$ 作为新的步长 $d_0(0 < \beta < 1)$,β 称为收缩因子, 一般取 $\beta \approx 0.4$. 如此重复迭代下去, 直到搜索步长小于给定正数 δ, 或者两次接连试验中的目标函数值之差小于给定正数 ε, 即

$$|f_k - f_{k+1}| < \varepsilon$$

迭代中止, 取 $\overline{x} = (x_k + x_{k+1})/2$ 作为近似极小值点.

一般说来, 用进退法求函数极小值点的效率较低, 实用意义不大. 但若把它稍加改进, 可用于寻找包含极小值点的区间. 前面介绍的 0.618 法, 斐波那契法和抛

物线法, 都是在包含一个极小值点的区间之内寻求 x^* 的方法, 因此用进退法确定这样的区间为这些更有效的一维搜索作好必要的准备. 其步骤如下:

设进行第 k 次试验时的初始点和步长分别为 x_{k-1} 和 d_{k-1}(d_{k-1} 可大于 0 或小于 0), 并且第 k 次试验成功, 即 $f_k < f_{k-1}$. 然后以 $x_k = x_{k-1} + d_{k-1}$ 作为第 $k+1$ 次试验的初始点, 以 $d_k = \alpha d_{k-1}$ 作为步长. 假定该次试验失败, 即 $f_k < f_{k+1} = f(x_{k+1}), x_{k+1} = x_k + d_k$. 显然这时有

$$\begin{cases} x_{k-1} < x_k < x_{k+1}(\text{或} x_{k-1} > x_k > x_{k+1}), \\ f_{k-1} > f_k < f_{k+1}, \end{cases} \tag{15.2.20}$$

因而区间 $[x_{k-1}, x_{k+1}]$ 内一定含有极小值点. 简而言之, 若一次试验成功之后紧接着一次试验失败, 由此确定的变量 x 的三个点构成的区间一定包含极小值点. 这三

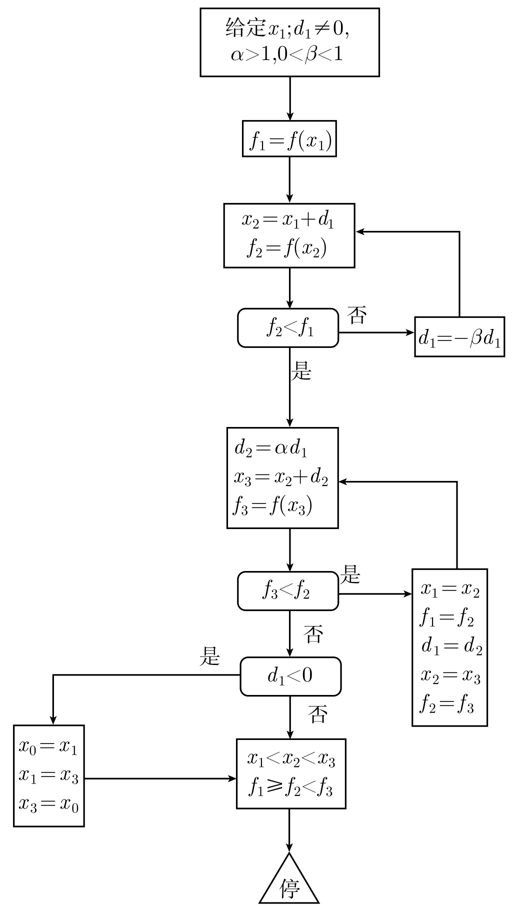


图 15.5　进退法确定包含极小值点区间的程序框图

个点可以作为抛物线内插的三个初始点. 进退法与抛物线法的这种组合, 对于一般的目标函数而言可能是普遍适用的最有效的一维搜索方法. 图 15.5 给出进退法确定搜索区间及三个初始点 x_1, x_2, x_3 的程序框图.

15.3 无约束 n 维极值的解析方法

无约束 n 维极值的计算方法所讨论的问题是

$$\min f(\boldsymbol{x}), \qquad \boldsymbol{x} \in E^n. \tag{15.3.1}$$

解法可以分为解析方法和直接方法两类. 利用目标函数一阶和二阶导数的方法称为解析方法. 因为函数的导数描述函数的变化规律, 所以利用导数的信息可加速向极小值点收敛的速度. 但在实际问题中, 有时目标函数难以给出导数的解析形式, 甚至目标函数本身就给不出解析表达式, 而只是知道若干特定变量值对应的目标函数值. 这时可以只利用函数值来求极小值点, 称为极小化的直接方法; 或者利用数值方法估计导数, 然后用解析法求解.

数值方法估计导数即利用差分作为导数的近似. 函数 $f(x)$ 的梯度 $g(x)$ 的第 i 个分量近似地等于

$$g_i = \frac{\partial f(\boldsymbol{x})}{\partial x_i} \approx \frac{f(x_1, \cdots, x_i + \Delta x_i, \cdots, x_n) - f(x_1, \cdots, x_i, \cdots, x_n)}{\Delta x_i}$$

或

$$g_i = \frac{\partial f(\boldsymbol{x})}{\partial x_i} \approx \frac{f(x_1, \cdots, x_i, \cdots, x_n) - f(x_1, \cdots, x_i - \Delta x_i, \cdots, x_n)}{\Delta x_i},$$
$$i = 1, 2, \cdots, n. \tag{15.3.2}$$

Δx_i 的有限大小导致的误差

$$\delta \approx \frac{\Delta x_i}{2} \frac{\partial^2 f(\boldsymbol{x})}{\partial x_i^2}\bigg|_{x=(x_1, \cdots, x_i, \cdots, x_n)}.$$

Δx_i 应选得足够小以减小误差, 但同时应使 δ 不小于计算机计算 f 值的舍入误差 (由计算机的有限字长所引起). 在二阶导数未知或不进行计算的情形下, Δx_i 的选择通常靠估计和猜测. 估计一阶导数比较保险的算式是

$$g_i = \frac{\partial f(\boldsymbol{x})}{\partial x_i} \approx \frac{f(x_1, \cdots, x_i + \Delta x_i, \cdots, x_n) - f(x_1, \cdots, x_i - \Delta x_i, \cdots, x_n)}{2\Delta x_i},$$
$$i = 1, 2, \cdots, n. \tag{15.3.3}$$

这样计算得到的 g_i 值误差大致正比于 $f(\boldsymbol{x})$ 的三阶导数, 因而误差较小; 但式 (15.3.3) 需要 $2n$ 个函数值, 而式 (15.3.2) 计算 g_i 只需 $(n+1)$ 个函数值, 比较省时.

二阶导数的计算更为费时, 幸运的是, 如果利用对称方法式 (15.3.3) 计算 g_i, 则二阶导数可借用其中的函数值, 因为二阶导数矩阵的对角元素可写成

$$G_{ii}=\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i^2}\approx[f(x_1,\cdots,x_i+\Delta x_i,\cdots,x_n)-f(x_1,\cdots,x_i-\Delta x_i,\cdots,x_n)$$
$$-2f(\boldsymbol{x})]/(\Delta x_i)^2,\qquad i=1,2,\cdots,n. \tag{15.3.4}$$

对非对角元素, 则有

$$G_{ij}=\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i\partial x_j}\approx[f(x_i+\Delta x_i,x_j+\Delta x_j)+f(x_i-\Delta x_i,x_j-\Delta x_j)$$
$$-f(x_i+\Delta x_i,x_j-\Delta x_j)-f(x_i-\Delta x_i,x_j+\Delta x_j)]/4\Delta x_i\Delta x_j,$$
$$i\neq j,i,\quad j=1,2,\cdots,n. \tag{15.3.5}$$

注意, 其中已使用了简化的记号. 由于二阶导数矩阵为对称矩阵, 即 $G_{ij}=G_{ji}$, 故共有 $n(n-1)/2$ 个独立的非对角线元素, 由式 (15.3.5) 可知, 需计算 $2n(n-1)$ 个目标函数值 (加上计算 g_i 时的 $2n$ 个目标函数值) 才能得到二阶导数矩阵 $\underset{\sim}{\boldsymbol{G}}$ 的完整表式.

如果目标函数的二阶导数在 $\boldsymbol{x}$ 附近的小区域内可近似地视为常数, 可以用下式代替式 (15.3.5) 估计非对角元素：

$$G_{ij}=\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i\partial x_j}$$
$$\approx\frac{f(x_i+\Delta x_i,x_j+\Delta x_j)+f(x_i,x_j)-f(x_i+\Delta x_i,x_j)-f(x_i,x_j+\Delta x_j)}{\Delta x_i\cdot\Delta x_j},$$
$$i\neq j,\qquad i,j=1,2,\cdots,n. \tag{15.3.6}$$

这样, 只需要 $n(n-1)/2$ 个函数值就能得到全部非对角元素 G_{ij}.

本节将介绍几种 n 维极值的解析方法, 直接方法留在 15.4 节叙述.

15.3.1　最速下降法 (梯度法)

设已有目标函数 $f(\boldsymbol{x})$ 的某个初始点 $\boldsymbol{x}_0$, 与一维极值算法一样, 我们要求产生一系列迭代点 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots$ 向极小值点 $\boldsymbol{x}^*$ 收敛, 也就是希望从迭代点 $\boldsymbol{x}_k$ 出发, 利用某种算法规则找到新的迭代点 $\boldsymbol{x}_{k+1}$, 它比 $\boldsymbol{x}_k$ 是 $\boldsymbol{x}^*$ 的更好近似. 为此, 将目标函数 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}_k$ 附近作泰勒展开

$$f(\boldsymbol{x})=f(\boldsymbol{x}_k)+(\boldsymbol{x}-\boldsymbol{x}_k)^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)+0\left(\|\boldsymbol{x}-\boldsymbol{x}_k\|\right), \tag{15.3.7}$$

其中最后一项表示 $\|\boldsymbol{x}-\boldsymbol{x}_k\|$ 的高阶小量. 当 $\|\boldsymbol{x}_{k+1}-\boldsymbol{x}_k\|$ 足够小并且

$$(\boldsymbol{x}_{k+1}-\boldsymbol{x}_k)^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)<\boldsymbol{0}$$

时, 有

$$f(\boldsymbol{x}_{k+1}) < f(\boldsymbol{x}_k),$$

即自变量 $\boldsymbol{x}$ 从 $\boldsymbol{x}_k$ 变到 $\boldsymbol{x}_{k+1}$ 时函数是下降的. 取 $(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k)$ 与负梯度方向一致

$$\boldsymbol{x}_{k+1} - x_k \propto -\nabla f(\boldsymbol{x}_k),$$

则 $-(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k)^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)$ 达到极大, 这是函数值下降最迅速的方向. 一个自然的想法是沿着函数值下降最迅速的方向进行极小值点的搜索是比较有效的, 根据这一原理可构造最迅速下降法的迭代方法如下：设已有了迭代点 $\boldsymbol{x}_k$, 则构造下一迭代点 $\boldsymbol{x}_{k+1}$ 的迭代公式为

$$\boldsymbol{x}_{k+1} - \boldsymbol{x}_k = -\lambda_k \cdot \nabla f(\boldsymbol{x}_k),$$

或写成

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \lambda_k \cdot \nabla f(\boldsymbol{x}_k), \tag{15.3.8}$$

其中 λ_k 是一正常数, 决定 $\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|$ 的长度, 称为**迭代步长**; $\boldsymbol{p}_k \equiv -\nabla f(\boldsymbol{x}_k)$ 称为搜索方向, 故式 (15.3.8) 亦可用 λ_k, $\boldsymbol{p}_k$ 表述为

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{p}_k. \tag{15.3.9}$$

式 (15.3.8) 的又一种表达形式是

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha_k \frac{\nabla f(\boldsymbol{x}_k)}{\|\nabla f(\boldsymbol{x}_k)\|} \equiv \boldsymbol{x}_k - \alpha_k \boldsymbol{s}_k, \tag{15.3.10}$$

其中 $\boldsymbol{s}_k$ 是模为 1 的单位矢量, 方向为目标函数在 $\boldsymbol{x}_k$ 点的梯度方向, 而

$$\alpha_k = \lambda_k \|\nabla f(\boldsymbol{x}_k)\| > 0.$$

迭代过程的终止可用如下的判据之一：

$$\begin{aligned} &|f(\boldsymbol{x}_{k+1}) - f(\boldsymbol{x}_k)| < \varepsilon_1, \\ &\left|\frac{\partial f(\boldsymbol{x}_k)}{\partial x_i}\right| < \varepsilon_2, \qquad i = 1, 2, \cdots, n, \\ &\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\| < \varepsilon_3. \end{aligned} \tag{15.3.11}$$

$\varepsilon_1, \varepsilon_2, \varepsilon_3$ 为给定的正数, 它们决定了近似极小值点的精确程度.

迭代法的收敛速度用量 γ 表示. 设函数极小值点为 $\boldsymbol{x}^*$, 第 k 和 $k+1$ 次迭代点分别是 $\boldsymbol{x}_k$ 和 $\boldsymbol{x}_{k+1}$, γ 定义为

$$\gamma = \lim_{k \to \infty} \frac{\|\boldsymbol{x}^* - \boldsymbol{x}_{k+1}\|}{\|\boldsymbol{x}^* - \boldsymbol{x}_k\|^p}. \tag{15.3.12}$$

$p=1$ 称为线性收敛, $p=n$ 称为 n 次收敛. p 值越大, 点列 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots$ 向 $\boldsymbol{x}^*$ 收敛速度越快.

现在求最速下降法的收敛速度. 将 $\nabla f(\boldsymbol{x}+\Delta\boldsymbol{x})$ 在 $\boldsymbol{x}$ 点展开为泰勒级数并取一次近似项

$$\nabla f(\boldsymbol{x}+\Delta\boldsymbol{x})=\nabla f(\boldsymbol{x})+\Delta\boldsymbol{x}\cdot\nabla^2 f(\boldsymbol{x}),$$

在极小值点附近有 $\Delta\boldsymbol{x}=\boldsymbol{x}^*-\boldsymbol{x}_k$, 即 $\boldsymbol{x}^*=\boldsymbol{x}_k+\Delta\boldsymbol{x}$, 于是

$$\nabla f(\boldsymbol{x}^*)=\nabla f(\boldsymbol{x}_k)+(\boldsymbol{x}^*-\boldsymbol{x}_k)\nabla^2 f(\boldsymbol{x}_k)=0.$$

据式 (15.3.10), 得

$$\begin{aligned}\boldsymbol{x}^*-\boldsymbol{x}_{k+1}&=\boldsymbol{x}^*-\boldsymbol{x}_k+\alpha_k\frac{\nabla f(\boldsymbol{x}_k)}{\|\nabla f(\boldsymbol{x}_k)\|}\\&=(\boldsymbol{x}^*-\boldsymbol{x}_k)+\frac{\alpha_k(\boldsymbol{x}_k-\boldsymbol{x}^*)\nabla^2 f(\boldsymbol{x}_k)}{\|\nabla f(\boldsymbol{x}_k)\|}\\&=(\boldsymbol{x}^*-\boldsymbol{x}_k)\left[\underset{\sim}{I}-\alpha_k\frac{\nabla^2 f(\boldsymbol{x}_k)}{\|\nabla f(\boldsymbol{x}_k)\|}\right].\end{aligned}$$

代入式 (15.3.12) 可知, $p=1$, 即最速下降法是收敛速度较慢的线性收敛.

现在回过来讨论迭代公式 (15.3.8) 或式 (15.3.9) 中步长 λ_k 的选取方法.

最简单的方法是每次迭代中步长为常数 λ, 这称为**定步长梯度法**. 在这种情况下可以证明, 存在一个与步长 λ 和目标函数 $f(\boldsymbol{x})$ 有关的正数 ε, 当 $\|\boldsymbol{x}_k-\boldsymbol{x}^*\|>\varepsilon$, 一定有 $f(\boldsymbol{x}_k)>f(\boldsymbol{x}_{k+1})$, 即每次迭代中目标函数值是下降的, $\boldsymbol{x}_{k+1}$ 较之 $\boldsymbol{x}_k$ 是 $\boldsymbol{x}^*$ 的更好近似. 但当 $\|\boldsymbol{x}_k-\boldsymbol{x}^*\|<\varepsilon$ 时, 只能保证

$$\|\boldsymbol{x}_r-\boldsymbol{x}^*\|\leqslant\varepsilon,\qquad r\geqslant k,$$

即继续迭代下去并不能进一步逼近极小值点 $\boldsymbol{x}^*$, 而是在 $\boldsymbol{x}^*\pm\varepsilon$ 的区间内围绕 $\boldsymbol{x}^*$ 来回振荡. 这就是说, 在定步长梯度法中, 步长一经给定, $\boldsymbol{x}^*$ 的近似值点的精确度也就确定了. 作为一种简单的补救办法, 在远离极小值点时, 步长可取得大一些, 以加速收敛; 当接近极小值点时步长应当减小, 以提高结果的精度.

由定步长梯度法得到一个启示, 迭代过程中应当取不同的步长值, 这种方案称为**变步长梯度法**. 如果每次迭代中使目标函数达到沿搜索方向的最小值点, 显然能加快计算速度, 这样的步长称为**最优步长**. 因此, **最优梯度法**的迭代公式与式 (15.3.8) 有相同的形式, 但其中的 λ_k 是一维函数 $\varphi(\lambda)\equiv f[\boldsymbol{x}_k-\lambda\nabla f(\boldsymbol{x}_k)]$ 的极小值点, 即

$$\varphi'(\lambda_k)=0. \tag{15.3.13}$$

这是一个一维搜索问题, 可用任一种一维搜索方法求解, 这里使用解析方法. 将 $f[\boldsymbol{x}_k-\lambda\nabla f(\boldsymbol{x}_k)]$ 在 $\boldsymbol{x}_k$ 的邻域作泰勒展开, 根据式 (15.1.5), 得

$$
\begin{aligned}
f(\boldsymbol{x}_k-\lambda\nabla f(\boldsymbol{x}_k)) = & f(\boldsymbol{x}_k)-\lambda\left[\nabla f(\boldsymbol{x}_k)\right]^{\mathrm{T}}\nabla f(\boldsymbol{x}_k) \\
& +\frac{\lambda^2}{2}\left[\nabla f(\boldsymbol{x}_k)\right]^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k)\nabla f(\boldsymbol{x}_k),
\end{aligned}
$$

将 $f\left[\boldsymbol{x}_k-\lambda\nabla f(\boldsymbol{x}_k)\right]$ 对 λ 求导数并令其等于 0, 求得 λ 的解 λ_k

$$
-\left[\nabla f(\boldsymbol{x}_k)\right]^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)+\lambda_k\left[\nabla f(\boldsymbol{x}_k)\right]^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k)\nabla f(\boldsymbol{x}_k)=0,
$$

$$
\lambda_k=\frac{\left[\nabla f(\boldsymbol{x}_k)\right]^{\mathrm{T}}\cdot\nabla f(\boldsymbol{x}_k)}{\left[\nabla f(\boldsymbol{x}_k)\right]^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k)\nabla f(\boldsymbol{x}_k)}. \tag{15.3.14}
$$

λ_k 即为最优步长.

由于负梯度方向具有最速下降性质, 一般容易认为这是理想的搜索方向. 实际上 $\boldsymbol{x}_k$ 点的负梯度方向仅在 $\boldsymbol{x}_k$ 点附近才具有最速下降性质, 是局部最优而不是全局最优. 最优步长法的实际使用表明, 开始几次迭代中, 步长比较大, 自变量的改变和函数值的下降也比较明显, 即收敛速度比较快. 但当接近极小点时步长很小, 目标函数值下降很慢, 即收敛速度很慢. 特别是, 最优梯度法是沿着锯齿形的路径向极小值点逼近的, 这一点可证明如下：由 $\varphi(\lambda)\equiv f\left[\boldsymbol{x}_k-\lambda\nabla f(\boldsymbol{x}_k)\right]$ 对 λ 求导, 得

$$
\begin{aligned}
\frac{\partial\varphi(\lambda)}{\partial\lambda} & =\frac{\partial f(\boldsymbol{x}_k-\lambda\nabla f(\boldsymbol{x}_k))}{\partial(\boldsymbol{x}_k-\lambda\nabla f(\boldsymbol{x}_k))}\cdot\frac{\partial(\boldsymbol{x}_k-\lambda\nabla f(\boldsymbol{x}_k))}{\partial\lambda} \\
& =-\left\{\nabla f\left[\boldsymbol{x}_k-\lambda\nabla f(\boldsymbol{x}_k)\right]\right\}^{\mathrm{T}}\nabla f(\boldsymbol{x}_k),
\end{aligned}
$$

由 $\varphi'(\lambda_k)=0$ 并注意 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k-\lambda_k\nabla f(\boldsymbol{x}_k)$, 得

$$
\left[\nabla f(\boldsymbol{x}_{k+1})\right]^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)=0. \tag{15.3.15}
$$

可见, 在相继的两个迭代点上, 函数 $f(\boldsymbol{x})$ 的两个梯度方向 $\nabla f(\boldsymbol{x}_k)$ 和 $\nabla f(\boldsymbol{x}_{k+1})$ 是相互正交的. 例如, 用最优梯度法搜索一个二维目标函数极小值点的过程如图 15.6 中的锯齿线所示, 图中的曲线表示目标函数的等值线. 由于梯度是函数的局部性质, 局部看来在一点附近目标函数下降虽然快了, 但从整体来看反而走了弯路.

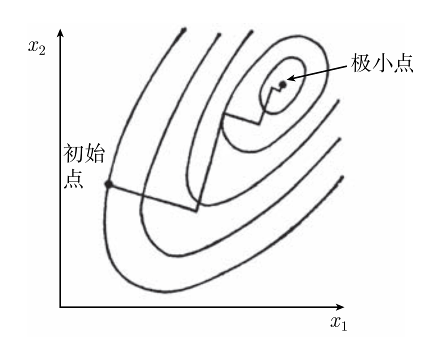


图 15.6　最优梯度法搜索二维函数极小值点的图示

梯度法虽然收敛速度较慢, 但这只在极小值点附近才比较明显. 它的优点是迭代过程简单, 一次迭代计算量小, 而且从远离极小值点 $\boldsymbol{x}^*$ 处的初始点开始, 迭代也能收敛到极小值点, 即对迭代初始点要求不严, 因此, 利用它来求得 $\boldsymbol{x}^*$ 的近似解还是合适的. 可以在该近似解的基础上再使用收敛速度更快的其他方法, 如 15.3.2 节介绍的牛顿法, 来求得 $\boldsymbol{x}^*$ 的更好的近似解.

15.3.2 牛顿法

如 15.1 节引言所述, 求函数极值问题可化为求解 $\nabla f(\boldsymbol{x}^*) = 0$ 的问题, 这是 n 个变量的非线性方程组. 求解非线性方程组的最古老方法是牛顿法, 它不但利用了函数在搜索点的梯度, 而且利用了它的二阶导数, 即考虑了梯度变化的趋势. 因此, 牛顿法的每一步搜索方向比最优梯度法有所改进.

首先以一维搜索为例. 假定已给出极小值点 x^* 的较好近似 x_k, 因为连续可导函数在极小值点附近的行为与二次函数很接近, 所以在 x_k 附近可用二次函数来逼近目标函数 $f(x)$, 即在点 x_k 对 $f(x)$ 作泰勒展开并取到二次项,

$$f(x) \approx \varphi(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}f''(x_k)(x - x_k)^2.$$

然后以二次函数 $\varphi(x)$ 的极小值点作为 $f(x)$ 极小值点的新的近似值 x_{k+1}. 由极值的必要条件 $\varphi'(x_{k+1}) = 0$ 可得

$$f'(x_k) + f''(x_k)(x_{k+1} - x_k) = 0.$$

由此求出一维搜索牛顿法的迭代公式

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}. \tag{15.3.16}$$

上述步骤可直接推广到 n 维极值的情形, 以上两式变为

$$\nabla f(\boldsymbol{x}_k) + \nabla^2 f(\boldsymbol{x}_k)(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k) = 0,$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \left[\nabla^2 f(\boldsymbol{x}_k)\right]^{-1} \nabla f(\boldsymbol{x}_k). \tag{15.3.17}$$

牛顿法中止迭代的判据可选为

$$\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\| < \varepsilon.$$

当目标函数 $f(\boldsymbol{x})$ 是二次函数, $f(\boldsymbol{x}) \equiv \varphi(\boldsymbol{x})$, 牛顿法求出的极值点解是严格正确的, 不管初始点选得怎样, 只需一次迭代就可求得极小值点. 对于非二次函数, 由于在极小值点附近它的行为与二次函数很接近, 故牛顿法的收敛速度也很快. 可以证明, 对于牛顿法有

$$\|\boldsymbol{x}^* - \boldsymbol{x}_{k+1}\| \sim \|\boldsymbol{x}^* - \boldsymbol{x}_k\|^2$$

即牛顿法是二次收敛的.

牛顿法在每次迭代中要计算目标函数 $f(\boldsymbol{x})$ 的二阶导数 (黑塞) 矩阵并求逆, 因而黑塞矩阵必须是非奇异的, 否则就不能用牛顿法计算. 为了保证迭代过程收敛, 黑塞矩阵的逆阵必须正定. 计算黑塞矩阵要计算 $\dfrac{1}{2}n(n+1)$ 个二阶偏导数, 对于高维的目标函数工作量很大, 占用计算机内存较多; 求黑塞矩阵的逆矩阵也是既困难又费时的问题. 牛顿法的另一缺点是, 当目标函数不是二次函数时, 要求初始点离极小值点不能太远, 否则不能保证迭代有好的收敛性, 或可能收敛向非极小值点, 但在极小值点未知的情形下要做到这一点是困难的. 如 15.3.1 节所述, 可以首先使用最速下降法找到一个较好的极小值点的近似作为牛顿法的初始值, 以此克服这一困难. 克服对初始点要求苛刻的另一种途径是利用**阻尼牛顿法**. 令

$$\boldsymbol{p}_k = -\left[\nabla^2 f(\boldsymbol{x}_k)\right]^{-1} \nabla f(\boldsymbol{x}_k), \tag{15.3.18}$$

代入牛顿法迭代公式 (15.3.17) 并与最速下降法迭代公式 (15.3.9) 做比较可知, 牛顿法相当于搜索方向 $\boldsymbol{p}_k$(由式 (15.3.18) 定义), 定步长 $\lambda_k = 1$ 的最速下降法. 这里 $\boldsymbol{p}_k$ 和 λ_k 分别称为牛顿搜索方向和牛顿步长. 在阻尼牛顿法中, 是将定步长改变为最优步长, 因而它的迭代公式可写成

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{p}_k, \tag{15.3.19}$$

其中 λ_k 满足

$$f(\boldsymbol{x}_k + \lambda_k \boldsymbol{p}_k) = \min_{\lambda} f(\boldsymbol{x}_k + \lambda \boldsymbol{p}_k). \tag{15.3.20}$$

λ_k 可用一维搜索方法或式 (15.3.14) 求出. 阻尼牛顿法保持了牛顿法的二次收敛特性, 对初始点又没有过于苛刻的要求, 但需要计算黑塞矩阵的逆阵的要求仍然存在.

15.3.3 共轭方向法和共轭梯度法

当目标函数是 n 维二次函数时, 应用**共轭方向法**可以在最多 n 次一维搜索内找到极小值点, 这种性质称为二次截止性质 (注意与二次收敛相区别). 对于非二次函数, 只要它具有连续二阶导数, 在极小值点附近目标函数可用二次函数作为近似, 因而**共轭方向**法也相当有效. 它克服了梯度法收敛慢、牛顿法计算量大的缺点.

首先介绍**共轭方向**的概念. 设 $\underset{\sim}{H}$ 为 $n \times n$ 对称正定矩阵, 如任意 n 维向量 $\boldsymbol{p}_i, \boldsymbol{p}_j$ 满足

$$\begin{aligned} &\boldsymbol{p}_i^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{p}_j = \boldsymbol{0}, \qquad i \neq j, \\ &\boldsymbol{p}_i^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{p}_i \neq \boldsymbol{0}, \end{aligned} \tag{15.3.21}$$

则称向量 $\boldsymbol{p}_i \boldsymbol{p}_j$ 是 $\underset{\sim}{H}$–共轭的. 如果取 $\underset{\sim}{H} = \underset{\sim}{I}_n$(单位矩阵), 则有

$$\boldsymbol{p}_i^{\mathrm{T}} \boldsymbol{p}_j = \boldsymbol{0} \ \ (i \neq j); \qquad \boldsymbol{p}_i^{\mathrm{T}} \boldsymbol{p}_i \neq \boldsymbol{0}.$$

即 $\boldsymbol{p}_i, \boldsymbol{p}_j (i \neq j)$ 是相互正交的向量. 所以 $\underset{\sim}{H}$–共轭是正交概念的推广, 正交是 $\underset{\sim}{H}$–共轭的特殊情形.

现在对共轭向量的性质稍加讨论. 设有–组 m 个 n 维向量 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_m$ 彼此 $\underset{\sim}{H}$–共轭, 则 $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_m$ 一定线性无关, 证明如下：如有一组数 $\alpha_1, \cdots, \alpha_m$ 满足

$$\alpha_1 \boldsymbol{p}_1 + \cdots + \alpha_m \boldsymbol{p}_m = 0,$$

则一定有

$$\boldsymbol{p}_i^{\mathrm{T}} \underset{\sim}{\boldsymbol{H}} (\alpha_1 \boldsymbol{p}_1 + \cdots + \alpha_m \boldsymbol{p}_m) = \boldsymbol{p}_i^{\mathrm{T}} \underset{\sim}{\boldsymbol{H}}\, \boldsymbol{0} = \boldsymbol{0}, \qquad i = 1, \cdots, m,$$

即

$$\sum_j \alpha_j \boldsymbol{p}_i^{\mathrm{T}} \underset{\sim}{\boldsymbol{H}}\, \boldsymbol{p}_j = \boldsymbol{0}, \qquad i, j = 1, 2, \cdots, m.$$

将式 (15.3.21) 代入得

$$\alpha_i = 0, \qquad i = 1, 2, \cdots, m.$$

所以 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_m$ 线性无关.

由于 n 维向量组最大的线性无关的向量个数是 n, 所以 n 维向量空间最多可以找出 n 个彼此 $\underset{\sim}{H}$–共轭的向量, 这 n 个向量构成 n 维空间的一个基.

很容易由一组 m 个线性无关的 n 维向量 $\boldsymbol{d}_1, \cdots, \boldsymbol{d}_m$ 来构造一组 m 个彼此 $\underset{\sim}{H}$–共轭的向量 $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_m$. 令 $\boldsymbol{p}_1 = \boldsymbol{d}_1$(实际上可令 $\boldsymbol{p}_i = \boldsymbol{d}_i, i$ 为 $1 \to m$ 中的任意指标, 这表示 $\boldsymbol{p}_1$ 的选择有很大的自由度), 容易证明

$$\boldsymbol{p}_2 = \boldsymbol{d}_2 + \alpha \boldsymbol{p}_1, \qquad \alpha = \frac{-\boldsymbol{p}_1^{\mathrm{T}} \underset{\sim}{H}\, \boldsymbol{d}_2}{\boldsymbol{p}_1^{\mathrm{T}} \underset{\sim}{H}\, \boldsymbol{p}_1}$$

与 $\boldsymbol{p}_1$ 是 $\underset{\sim}{H}$–共轭的. 事实上, $\boldsymbol{d}_2$ 可用 $\boldsymbol{d}_3, \cdots, \boldsymbol{d}_m$ 中任一个代替, 所以 $\boldsymbol{p}_1$ 的构成也有很大的自由度. 进一步, 设已求出 $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_k$ 是彼此 $\underset{\sim}{H}$–共轭的, 而且它们中的每一个都是 $\boldsymbol{d}_1, \cdots, \boldsymbol{d}_m$ 的线性组合, 通过简单的计算知道

$$\boldsymbol{p}_{k+1} = \boldsymbol{d}_{k+1} - \sum_{i=1}^{k} \frac{\boldsymbol{p}_i^{\mathrm{T}} \underset{\sim}{H}\, \boldsymbol{d}_{k+1}}{\boldsymbol{p}_i^{\mathrm{T}} \underset{\sim}{H}\, \boldsymbol{p}_i} \boldsymbol{p}_i \tag{15.3.22}$$

与 $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_k$ $\underset{\sim}{H}$–共轭. 由于 $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_k$ 是 $\boldsymbol{d}_1, \cdots, \boldsymbol{d}_m$ 的线性组合, 式 (15.3.22) 右边第二项中 $\boldsymbol{p}_1$ 的系数是一标量, 所以 $\boldsymbol{p}_{k+1}$ 也是 $\boldsymbol{d}_1, \cdots, \boldsymbol{d}_m$ 的线性组合. 如 $\boldsymbol{p}_1, \boldsymbol{p}_2$ 的构成所表明的, 在构成 $\boldsymbol{p}_i, i = 1, \cdots, m-1$ 时, $\boldsymbol{d}_i$ 的选择都有一定的自由度, 所以共轭向量 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_m$ 不是唯一的.

利用式 (15.3.22), 取 $k=1,\cdots,m-1$, 从一组 m 个线性无关的向量 $\boldsymbol{d}_1,\cdots,\boldsymbol{d}_m$ 构成 m 个彼此 $\underset{\sim}{H}$–共轭的向量 $\boldsymbol{p}_1,\cdots,\boldsymbol{p}_m$ 称为共轭化过程.

现在讨论二次函数的极小化问题. n 维的二次函数的一般形式可表为

$$f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\underset{\sim}{H}\boldsymbol{x}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}+c, \qquad \bar{x}\in E^n, \tag{15.3.23}$$

其中 $\underset{\sim}{H}$ 为 $n\times n$ 矩阵, $\boldsymbol{b}$ 为 n 维向量. 设 $f(\boldsymbol{x})$ 的极小值点为 $\boldsymbol{x}^*,\boldsymbol{x}_0$ 为任一给定的初始点. 若向量 $\boldsymbol{p}_0,\cdots,\boldsymbol{p}_{n-1}$ 彼此 $\underset{\sim}{H}$–共轭, 它们构成 n 维空间的一个基, 任何一个 n 维向量可唯一地表示成这组向量的线性组合. 因此, 向量 $\boldsymbol{x}^*-\boldsymbol{x}_0$ 可表示为

$$\boldsymbol{x}^*-\boldsymbol{x}_0=\alpha_0\boldsymbol{p}_0+\cdots+\alpha_{n-1}\boldsymbol{p}_{n-1},$$

或写成

$$\boldsymbol{x}^*=\boldsymbol{x}_0+\alpha_0\boldsymbol{p}_0+\cdots+\alpha_{n-1}\boldsymbol{p}_{n-1}, \tag{15.3.24}$$

这就将 $f(\boldsymbol{x})$ 的极小化问题转化为求 n 个数 $\alpha_0,\cdots,\alpha_{n-1}$ 的问题.

我们分析一下将彼此 $\underset{\sim}{H}$–共轭的 n 个方向作为基所带来的好处. 设 $\boldsymbol{x}_k$ 为第 k 次迭代点, 从 $\boldsymbol{x}_k$ 出发, 沿方向 $\boldsymbol{p}_k$ 求下一迭代点 $\boldsymbol{x}_{k+1}$, 按最优梯度法, 其最优步长 λ_k^* 应满足

$$f(\boldsymbol{x}_k+\lambda_k^*\boldsymbol{p}_k)=\min_{\lambda} f(\boldsymbol{x}_k+\lambda\boldsymbol{p}_k),$$

亦即

$$\left.\frac{\partial f(\boldsymbol{x}_k+\lambda\boldsymbol{p}_k)}{\partial\lambda}\right|_{\lambda=\lambda_k^*}=0.$$

记 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k^*\boldsymbol{p}_k$, 上式相当于

$$[\nabla f(\boldsymbol{x}_{k+1})]^{\mathrm{T}}\boldsymbol{p}_k=0. \tag{15.3.25}$$

即任一次迭代中的搜索方向 $\boldsymbol{p}_k$ 与下一迭代点的梯度方向正交. 这一结果事实上已在最优梯度法中导出 (见式 (15.3.15) 并注意 $\boldsymbol{p}_k=-\nabla f(\boldsymbol{x}_k)$).

对于式 (15.3.23) 所表示的二次函数, 有

$$\nabla f(\boldsymbol{x})=\underset{\sim}{H}\boldsymbol{x}+\boldsymbol{b}, \tag{15.3.26}$$

代入式 (15.3.25), 可求得

$$\lambda_k^*=\frac{-[\nabla f(\boldsymbol{x}_k)]^{\mathrm{T}}\boldsymbol{p}_k}{\boldsymbol{p}_k^{\mathrm{T}}\underset{\sim}{H}\boldsymbol{p}_k}. \tag{15.3.27}$$

由于 $\boldsymbol{x}^*$ 是 $f(\boldsymbol{x})$ 的极小值点, 故对二次函数应有

$$\nabla f(\boldsymbol{x}^*)=\underset{\sim}{H}\boldsymbol{x}^*+\boldsymbol{b}=\boldsymbol{0}.$$

将式 (15.3.24) 两边左乘 $\boldsymbol{p}_0^{\mathrm{T}} \underset{\sim}{H}$, 注意 $\boldsymbol{p}_0^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{p}_j = \mathbf{0}(j=1,\cdots,n-1)$, 得到

$$\begin{aligned}&\boldsymbol{p}_0^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{x}^* = \boldsymbol{p}_0^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{x}_0 + \alpha_0 \boldsymbol{p}_0^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{p}_0,\\&\boldsymbol{p}_0^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{x}^* + \boldsymbol{p}_0^{\mathrm{T}} \boldsymbol{b} - \boldsymbol{p}_0^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{x}_0 - \boldsymbol{p}_0^{\mathrm{T}} \boldsymbol{b} = \alpha_0 \boldsymbol{p}_0^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{p}_0,\end{aligned}$$

注意式 (15.3.26) 及 $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$, 则得

$$-\boldsymbol{p}_0^{\mathrm{T}} \nabla f(\boldsymbol{x}_0) = \alpha_0 \boldsymbol{p}_0^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{p}_0,$$

亦即

$$\alpha_0 = \frac{-\boldsymbol{p}_0^{\mathrm{T}} \nabla f(\boldsymbol{x}_0)}{p_0^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{p}_0}.$$

与式 (15.3.27) 对比可知, α_0 就是从 $\boldsymbol{x}_0$ 出发沿共轭方向 $\boldsymbol{p}_0$ 求二次函数 $f(\boldsymbol{x})$ 的极小值点的最优步长 λ_0^*.

记

$$\boldsymbol{x}_k = \boldsymbol{x}_0 + \alpha_0 \boldsymbol{p}_0 + \cdots + \alpha_{k-1} \boldsymbol{p}_{k-1},$$

于是

$$\boldsymbol{x}^* = \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k + \cdots + \alpha_{n-1} \boldsymbol{p}_{n-1},$$

两边左乘 $\boldsymbol{p}_k^{\mathrm{T}} \underset{\sim}{H}$, 注意 $\boldsymbol{p}_k^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{p}_j = \mathbf{0}\ (j=k+1,\cdots,n-1)$, 得到

$$\boldsymbol{p}_k^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{x}^* = \boldsymbol{p}_k^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{p}_k,$$

做类似于上面的推导, 得

$$\alpha_k = \frac{-\boldsymbol{p}_k^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)}{\boldsymbol{p}_k^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{p}_k}, \qquad k = 0,1,\cdots,n-1. \tag{15.3.28}$$

与式 (15.3.27) 比较可知, α_k 就是从迭代点 $\boldsymbol{x}_k$ 出发沿共轭方向 $\boldsymbol{p}_k$ 求二次函数 $f(\boldsymbol{x})$ 的极小值点的最优步长 λ_k^*.

由此得出结论: 欲求式 (15.3.23) 所示的二次函数 $f(\boldsymbol{x})$ 的极小值点, 若有了彼此 $\underset{\sim}{H}$–共轭的 n 个方向 $\boldsymbol{p}_0, \boldsymbol{p}_1 \cdots, \boldsymbol{p}_{n-1}$, 可从任何初始点 $\boldsymbol{x}_0$ 出发, 分别沿这 n 个共轭方向做一维搜索, 求出 n 个最优步长即为系数 $\alpha_k(k=0,1,\cdots,n-1)$, 代入式 (15.3.24), 即求得极小值点 $\boldsymbol{x}^*$, 这就是共轭方向法的要点. 由于 α_k 可以为 0, 所以最多做 n 次一维搜索就可解得 $f(\boldsymbol{x})$ 的极小值点.

当 $f(\boldsymbol{x})$ 不是二次函数时, 在所有迭代点 $\boldsymbol{x}_k$, 利用二次函数 $\varphi_k(\boldsymbol{x})$($f(\boldsymbol{x})$ 的泰勒展开取到二次项) 作为 $f(\boldsymbol{x})$ 的近似

$$f(\boldsymbol{x}) \approx \varphi_k(\boldsymbol{x}) = f(\boldsymbol{x}_k) + (\boldsymbol{x} - \boldsymbol{x}_k)^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)$$

$$+\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_k)^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k)(\boldsymbol{x}-\boldsymbol{x}_k),$$

求 $\varphi_k(\boldsymbol{x})$ 的极小值作为新的迭代点, 这一过程可用共轭方向法完成, 与式 (15.3.23) 对比知, 这时式中的 $\underset{\sim}{H}$ 就成为目标函数的黑塞矩阵. 每迭代一个点便依上式更换一个二次函数, 直到满足迭代中止判据

$$\|\boldsymbol{x}_{k+1}-\boldsymbol{x}_k\|<\varepsilon \quad 或 \quad \|\nabla f(\boldsymbol{x}_{k+1})\|<\varepsilon$$

为止, $\boldsymbol{x}_{k+1}$ 取为近似极小值点.

前面已述, 共轭方向的选取有很大的任意性, 用不同方式产生共轭方向就得到共轭方向法的不同方案. 产生一组共轭方向的最简单方法是利用 n 维坐标空间中的 n 个坐标单位向量 $\boldsymbol{e}_i(i=1,2,\cdots,n)$ 作为一组线性无关的向量

$$\begin{aligned}\boldsymbol{e}_1&=(1,0,0,\cdots,0,0),\\ \boldsymbol{e}_2&=(0,1,0,\cdots,0,0),\\ &\vdots\\ \boldsymbol{e}_n&=(0,0,0,\cdots,0,1).\end{aligned}$$

令 $\boldsymbol{d}_i=\boldsymbol{e}_i$, 代入式 (15.3.22) 构成一组 n 个共轭向量 $\boldsymbol{p}_1,\cdots,\boldsymbol{p}_n$. 这种方案的缺点是需要用到黑塞矩阵, 所以对于一般的目标函数, 每迭代一次要计算一次二阶导数矩阵, 这一缺点与牛顿法一样未得到克服.

我们可以只利用式 (15.3.23) 所示的目标函数 $f(\boldsymbol{x})$ 的梯度来构造一组共轭方向, 以避免黑塞矩阵的繁杂计算. 这种方案既用到目标函数梯度, 又用到共轭方向, 故称为**共轭梯度法**. 具有代表性的共轭梯度法是 1964 年弗莱彻和吕伍斯 (Fletcher, Reeves) 提出的方案.

设已求得二次函数 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}_0$ 和 $\boldsymbol{x}_1$ 点的梯度 $\boldsymbol{g}_0=\nabla f(\boldsymbol{x}_0)$ 和 $\boldsymbol{g}_1=\nabla f(\boldsymbol{x}_1)$, 由式 (15.3.26) 易得

$$\boldsymbol{g}_1-\boldsymbol{g}_0=\underset{\sim}{H}(\boldsymbol{x}_1-\boldsymbol{x}_0). \tag{15.3.29}$$

任何与 $\Delta\boldsymbol{g}=\boldsymbol{g}_1-\boldsymbol{g}_0$ 正交的矢量 $\boldsymbol{p}$ 与 $\Delta\boldsymbol{x}=\boldsymbol{x}_1-\boldsymbol{x}_0$ 是 $\underset{\sim}{H}$–共轭的, 因为

$$\boldsymbol{p}^{\mathrm{T}}(\boldsymbol{g}_1-\boldsymbol{g}_0)=\boldsymbol{p}^{\mathrm{T}}\underset{\sim}{H}(\boldsymbol{x}-\boldsymbol{x}_0)=0. \tag{15.3.30}$$

利用这一性质, 可以无须二阶导数矩阵 $\underset{\sim}{H}$ 的知识, 而依靠梯度的变化来得到共轭方向.

共轭梯度法的迭代过程可描述如下. 从任一初始点 $\boldsymbol{x}_0$ 出发, 第一次搜索方向是 $\boldsymbol{x}_0$ 点的最速下降方向即 $\boldsymbol{p}_0=-\boldsymbol{g}_0$. 沿该方向按最优梯度法找到的极小值点记为 $\boldsymbol{x}_1$, 即

$$\boldsymbol{x}_1-\boldsymbol{x}_0=\lambda_0\boldsymbol{p}_0. \tag{15.3.31}$$

在 $\boldsymbol{x}_1$ 点梯度记为 $\boldsymbol{g}_1$. 按共轭梯度法, 下一搜索方向 $\boldsymbol{p}_1$ 应与 $\boldsymbol{p}_0$ 彼此 $\underset{\sim}{H}$–共轭, $\boldsymbol{p}_1$ 可由 $\boldsymbol{p}_0$ 和 $\boldsymbol{g}_1$ 的线性组合构成

$$\boldsymbol{p}_1 = -\boldsymbol{g}_1 + \alpha_0 \boldsymbol{p}_0. \tag{15.3.32}$$

由共轭条件应有

$$\boldsymbol{p}_1^{\mathrm{T}} \underset{\sim}{H} \boldsymbol{p}_0 = \boldsymbol{0},$$

代入式 (15.3.31), 即为

$$\boldsymbol{p}_1^{\mathrm{T}} \underset{\sim}{H} (\boldsymbol{x}_1 - \boldsymbol{x}_0) = \boldsymbol{0}.$$

将式 (15.3.32) 代入并注意式 (15.3.29), 有

$$(-\boldsymbol{g}_1 + \alpha_0 \boldsymbol{p}_0)^{\mathrm{T}} \underset{\sim}{H} (\boldsymbol{x}_1 - \boldsymbol{x}_0) = (-\boldsymbol{g}_1 - \alpha_0 \boldsymbol{p}_0)^{\mathrm{T}} (\boldsymbol{g}_1 - \boldsymbol{g}_0) = 0.$$

由式 (15.3.25) 知

$$\boldsymbol{g}_{k+1}^{\mathrm{T}} \boldsymbol{p}_k = 0,$$

令 $k=0$ 并代入前式, 求得

$$\alpha_0 = \frac{\boldsymbol{g}_1^{\mathrm{T}} \boldsymbol{g}_1}{\boldsymbol{g}_0^{\mathrm{T}} \boldsymbol{g}_0}. \tag{15.3.33}$$

该式与式 (15.3.32) 一起构成了新的共轭方向 $\boldsymbol{p}_1$. 类似的推导可产生所有 n 个共轭方向, 一般性的公式是

$$\begin{aligned}
&\boldsymbol{p}_0 = -\boldsymbol{g}_0, \\
&\boldsymbol{p}_{k+1} = -\boldsymbol{g}_{k+1} + \alpha_k \boldsymbol{p}_k, \\
&\alpha_k = \frac{\boldsymbol{g}_{k+1}^{\mathrm{T}} \boldsymbol{g}_{k+1}}{\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{g}_k}, \qquad k = 0, 1, \cdots, n-2.
\end{aligned} \tag{15.3.34}$$

这样产生的 n 个方向中的任一个与其余 $n-1$ 个方向彼此都 $\underset{\sim}{H}$–共轭.

可以证明, 如果取 $\boldsymbol{p}_0 \neq -\boldsymbol{g}_0$, 对于一个二次函数的目标函数, 弗莱丘–吕伍斯共轭梯度法有时 n 步之内不能达到极小值点, 这时应将 n 次搜索后到达的变量点 $\boldsymbol{x}_n$ 作为新的初始值 $\boldsymbol{x}_0$, 用共轭梯度法进行新的一周期 n 次搜索, 才能收敛到极小值点, 否则有可能无休止地迭代下去.

对于非二次函数, 可用

$$\alpha_k = \frac{\boldsymbol{g}_{k+1}^{\mathrm{T}} \boldsymbol{g}_{k+1} - \boldsymbol{g}_{k+1}^{\mathrm{T}} \boldsymbol{g}_k}{\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{g}_k} \tag{15.3.35}$$

代替式 (15.3.34) 中的系数 α_k.

共轭梯度法因为也是共轭方向法, 所以也有二次截止的性质. 它的优点是不必计算二阶导数矩阵, 因而节省了计算机存储和计算工作量; 它的收敛速度虽不如牛顿法但比最优梯度法快, 计算量比梯度法大但小于牛顿法. 共轭梯度法的程序框图如图 15.7.

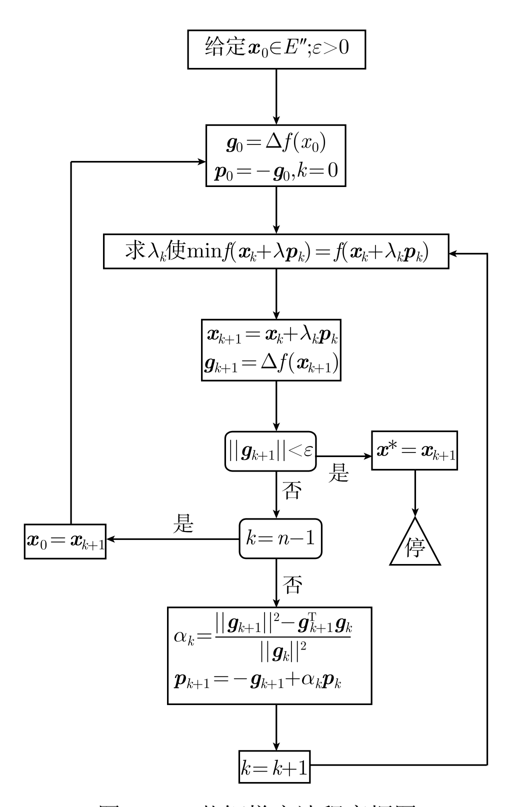


图 15.7　共轭梯度法程序框图

15.3.4　变尺度法

变尺度法在 1959 年由戴维登 (Davidon) 首先提出, 1963 年为弗莱彻和鲍威尔 (Powell) 所简化, 所以也称为 DFP 方法. 它与共轭梯度法一样, 是为了克服梯度法收敛慢、牛顿法计算量大的缺点而提出的.

记得牛顿法的迭代公式是

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{p}_k,$$

其中 $\boldsymbol{p}_k$ 是牛顿搜索方向, 它可表示为

$$\boldsymbol{p}_k=-\left[\nabla^2 f(\boldsymbol{x}_k)\right]^{-1}\nabla f(x_k)\equiv-\left[\underset{\sim}{H}(\boldsymbol{x}_k)\right]^{-1}\boldsymbol{g}(x_k).$$

变尺度法针对牛顿法要计算黑塞矩阵的逆矩阵 $\left[\underset{\sim}{H}(\boldsymbol{x}_k)\right]^{-1}$ 这个缺点, 利用一个 $n\times n$ 矩阵 $\underset{\sim k}{A}$ 来逼近 $\left[\underset{\sim}{H}(\boldsymbol{x}_k)\right]^{-1}$, 即取搜索方向为

$$\boldsymbol{p}_k = -\underset{\sim k}{A}^{\mathrm{T}}\boldsymbol{x}_k. \tag{15.3.36}$$

$\underset{\sim k}{A}$ 称为**方向矩阵**, 要求满足迭代关系

$$\underset{\sim k+1}{A} = \underset{\sim k}{A} + \Delta\underset{\sim k}{A}, \tag{15.3.37}$$

因此, 若能给出初始矩阵 $\underset{\sim 0}{A}$ 及 $\Delta\underset{\sim k}{A}$ 的表达式, 就建立了 $\underset{\sim k}{A}$ 的迭代公式.

这里我们不加证明地给出 DFP 方法中 $\Delta\underset{\sim k}{A}$ 的表达式如下：

$$\begin{aligned}
&\Delta\underset{\sim k}{A} = \underset{\sim k}{B} - \underset{\sim k}{C},\\
&\underset{\sim k}{B} = \frac{\Delta\boldsymbol{x}_k(\Delta\boldsymbol{x}_k)^{\mathrm{T}}}{(\Delta\boldsymbol{x}_k)^{\mathrm{T}}\Delta\boldsymbol{g}_k},\\
&\underset{\sim k}{C} = \frac{\underset{\sim k}{A}\Delta\boldsymbol{g}_k(\underset{\sim k}{A}\Delta\boldsymbol{g}_k)^{\mathrm{T}}}{\Delta\boldsymbol{g}_k^{\mathrm{T}}\underset{\sim k}{A}\Delta\boldsymbol{g}_k},
\end{aligned} \tag{15.3.38}$$

其中 $\Delta\boldsymbol{g}_k = \boldsymbol{g}_{k+1} - \boldsymbol{g}_k$, $\Delta\boldsymbol{x}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k$; $\boldsymbol{x}_{k+1}$ 是从 $\boldsymbol{x}_k$ 出发进行一维搜索确定最优步长 λ_k 后求得的下一迭代点, 即

$$\min_{\lambda} f(\boldsymbol{x}_k + \lambda\boldsymbol{p}_k) = f(\boldsymbol{x}_k + \lambda_k\boldsymbol{p}_k) = f(\boldsymbol{x}_{k+1}).$$

$\underset{\sim k}{B}$ 和 $\underset{\sim k}{C}$ 都是 $n\times n$ 维矩阵. 一般地 $\underset{\sim k}{A}$ 的初始矩阵 $\underset{\sim 0}{A}$ 取为单位矩阵 $\underset{\sim n}{I}$, 于是

$$\begin{aligned}
\underset{\sim k+1}{A} &= \underset{\sim k}{A} + \Delta\underset{\sim k}{A} = \underset{\sim k-1}{A} + \Delta\underset{\sim k-1}{A} + \Delta\underset{\sim k}{A} + \cdots\\
&= \underset{\sim 0}{A} + \sum_{i=0}^{k}\Delta\underset{\sim i}{A} = \underset{\sim n}{I} + \sum_{i=0}^{k}B_i - \sum_{i=0}^{k}C_i.
\end{aligned} \tag{15.3.39}$$

由以上所述, 得到 DFP 变尺度法的迭代步骤如下：

(1) 给定初始点 $\boldsymbol{x}_0 \in E^n$, 误差控制常数 ε.

(2) 令 $0 \Rightarrow k$, $\underset{\sim 0}{A} = \underset{\sim n}{I}$.

(3) 求 $\boldsymbol{g}_k = \nabla f(x_k)$, $\boldsymbol{p}_k = -\underset{\sim k}{A}\boldsymbol{g}_k$.

(4) 沿 $\boldsymbol{p}_k$ 进行一维搜索确定最优步长 λ_k 和下一迭代点 $\boldsymbol{x}_{k+1}$

$$\min_{\lambda} f(\boldsymbol{x}_k + \lambda\boldsymbol{p}_k) = f(\boldsymbol{x}_k + \lambda_k\boldsymbol{p}_k) = f(\boldsymbol{x}_{k+1});$$

(5) $\Delta \boldsymbol{x}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k$, $\boldsymbol{g}_{k+1} = \nabla f(\boldsymbol{x}_{k+1})$, $\Delta \boldsymbol{x}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k$, 若 $\|\Delta \boldsymbol{x}_k\| < \varepsilon$, 或 $\|\boldsymbol{g}_{k+1}\| < \varepsilon$, 则求出最优解 $\boldsymbol{x}^* \approx \boldsymbol{x}_{k+1}$, 迭代停止; 否则进行下一步.

(6) 若 $k = n$, 则 $\boldsymbol{x}_{k+1} \Rightarrow \boldsymbol{x}_0$, 转向 (2); 否则进行下一步.

(7) 按式 (15.3.38) 和式 (15.3.39) 求出 $\underset{\sim k+1}{A}$, 令 $k+1 \Rightarrow k$, 返回 (3).

可以证明, 变尺度法有以下性质:

(1) 若矩阵 $\underset{\sim k}{A}$ 为正定, 则 $\underset{\sim k+1}{A}$ 也为正定.

(2) 对于式 (15.3.23) 所表示的二次 (目标) 函数 $f(\boldsymbol{x})$ 有

(a) 在最优点方向矩阵 $\underset{\sim k}{A}$ 收敛于 $f(\boldsymbol{x})$ 的二阶导数矩阵的逆阵 $\left[\underset{\sim}{H}(\boldsymbol{x})\right]^{-1}$.

(b) 相连的两次迭代中的搜索方向是 $\underset{\sim}{H}$–共轭的, 即 DFP 法既是阻尼牛顿法, 又是共轭梯度法, 因此具有二次截止性质, 即在 n 次迭代内收敛到极小值点.

对于一般的目标函数, 可在每个迭代点用二次函数作为近似, 这时矩阵 $\underset{\sim k}{A}$ 亦能很快收敛.

由于变尺度法也是共轭梯度法, 如同 15.3.3 节所述, 在 n 次迭代后, 如果尚未收敛到极小值点, 应将变量点 $\boldsymbol{x}_n \Rightarrow \boldsymbol{x}_0$, 重新开始新的周期的迭代, 这就是迭代步骤 (6) 存在的原因.

变尺度法的收敛速度介于梯度法与牛顿法之间, 并且在一般情况下优于共轭梯度法. 它兼有共轭梯度法和牛顿法两者的优点, 是无约束极小化最通用的有效方法之一.

20 世纪 60 年代后期大量使用 DFP 方法发现, 对某些问题它的解在稳定性方面存在一些问题. 1970 年, 布鲁登 (Broyden), 弗莱彻 (Fletcher), 戈德史汀 (Goldstein), 香农 (Shannon) 等人导出了更稳定的算法, 简称 BFGS 变尺度法, 它的迭代公式为

$$\begin{aligned}\underset{\sim k+1}{A'} = \underset{\sim k+1}{A} + \frac{1}{(\Delta \boldsymbol{x}_k)^{\mathrm{T}} \Delta \boldsymbol{g}_k} \Bigg\{ & \frac{\boldsymbol{z}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k}{(\Delta \boldsymbol{x}_k)^{\mathrm{T}} \Delta \boldsymbol{g}_k} \Delta \boldsymbol{x}_k (\Delta \boldsymbol{x}_k)^{\mathrm{T}} \\ & - \Delta \boldsymbol{x}_k (\boldsymbol{z}_k)^{\mathrm{T}} - \boldsymbol{z}_k (\Delta \boldsymbol{x}_k)^{\mathrm{T}} \Bigg\} + \frac{\boldsymbol{z}_k (\boldsymbol{z}_k)^{\mathrm{T}}}{(\boldsymbol{z}_k)^{\mathrm{T}} \Delta \boldsymbol{g}_k},\end{aligned}$$

其中 $\boldsymbol{z}_k = \underset{\sim k}{A} \Delta \boldsymbol{g}_k$, $\boldsymbol{A}'_{k+1}$ 是 BFGS 变尺度法的迭代矩阵, $\underset{\sim k+1}{A}$ 为 DFP 变尺度法的迭代矩阵. BFGS 法用于高维目标函数有较好的稳定性.

15.4　无约束 n 维极值的直接方法

15.3 节描述的方法都用到目标函数 $f(\boldsymbol{x})$ 的一阶或二阶导数 $\nabla f(\boldsymbol{x})$, $\nabla^2 f(\boldsymbol{x})$. 但实标问题中的目标函数有时很复杂, 有的甚至没有明显的解析表达式, 只有函数值与变量的对应关系, 因而导数难以求得或根本不存在, 有时虽然可用数值方法求

出一阶、二阶导数, 但大大增加了计算工作量. 在这种情形下, 利用直接方法比较适宜. 直接方法的优点是对目标函数的解析性质没有太多的要求, 因而适用面较宽; 反过来正因为它不利用函数的解析性质, 难以判断函数的变化趋势, 所以收敛速度较慢, 计算量往往按变量维数 n 的幂次增加, 导致它在高维问题上应用的极大困难.

15.4.1　坐标轮换法

坐标轮换法又称**变量轮换法**, 这是一种最古老的多维搜索方案. 对于 n 维目标函数的极小值问题, 它的迭代过程是沿着 n 个不同的坐标方向轮换地进行一维搜索. 设初始点表示为 $\boldsymbol{x}_0=\left\{x_0^{(1)},x_0^{(2)},\cdots,x_0^{(n)}\right\}$, 首先对一个变量, 例如, $x_0^{(1)}$ 进行一维寻优而保持其余 $n-1$ 个变量值不变, 找到的第一个极小点令为 $\boldsymbol{x}_1$; 然后对第二个变量 $x_0^{(2)}$ 进行一维寻优而保持其余 $n-1$ 个变量值不变, 找到第二个极小点, 令为 $\boldsymbol{x}_2$; 依次对所有 n 个变量做类似的寻优, 得到第 n 个极小点为 $\boldsymbol{x}_n$, 这就完成了一次迭代. 若 $\|\boldsymbol{x}_n-\boldsymbol{x}_0\|<\varepsilon$, ε 为预先给定的误差控制常数, 则 $\boldsymbol{x}_n$ 取为近似极小值点, 迭代中止; 否则令 $\boldsymbol{x}_0=\boldsymbol{x}_n$, 重复以上迭代过程. 一维寻优可用 15.2 节中介绍的任何一种方法.

坐标轮换法的收敛速度与目标函数的形状有直接关系. 以二元函数为例, 有以下几种情况：

(1) 目标函数等值线为圆或长、短轴平行于坐标轴的椭圆, 则两次一维搜索即可达到极小值点. 这种情形如图 15.8(a) 所示, 图中曲线表示目标函数的等值线, 直线段表示一维寻优的搜索方向和步长.

(2) 目标函数等值线为长、短轴不平行于坐标轴的椭圆, 则要多于两次的一维搜索才能达到极小值点, 见图 15.8(b).

(3) 对于函数等值线如图 15.8(c) 所示的目标函数, 存在着与坐标轴不相平行的“脊线”, 本来沿着脊线方向寻优是比较好的搜索方向, 可由少数 n 步一维搜索即达到极小. 但坐标轮换法的一维搜索沿着平行于坐标轴方向进行, 因此收敛速度极慢.

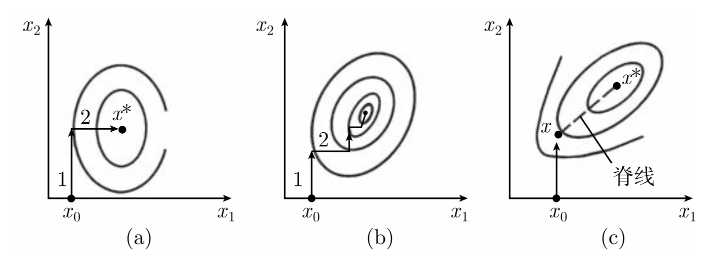


图 15.8　多维目标函数极小点搜索的坐标轮换法

对于 n 维目标函数 $f(x_1,x_2,\cdots,x_n)$, 如果共需 m 次迭代能收敛到极小点, 每次迭代有 n 个一维搜索, 假定每次一维搜索平均要计算 p 次目标函数值, 则达到极

小值点需计算的目标函数值次数 k 为 $k = mp^n$. 对于高维极小化问题, k 值往往非常大, 使得坐标轮换法实际上无法完成.

15.4.2　霍克–吉弗斯模式搜索法

该方法又称**步长加速法**, 是霍克–吉弗斯 (Hooke-Jeeves)1961 年提出的. 15.4.1 节中我们看到, 沿目标函数的脊线方向搜索是到达极小值点的有利方向. 本方法正是近似地沿脊线方向进行搜索的方法, 它的基本思想是, 进行试验性搜索后发现, 沿某个方向移动能使目标函数值减小, 则继续沿该方向移动可能是有效的.

模式搜索中包含两类 "移动": **探测移动**和**模式移动**. 前者是为了寻找目标函数的下降方向, 后者是沿着该方向寻找极小值点.

设目标函数为 $f(\boldsymbol{x}), \boldsymbol{x} \in E^n$. 给定初始点 $\boldsymbol{x}_0$ 和步长向量 $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}^{\mathrm{T}}$, 并设 $\boldsymbol{e}_i$ 为 n 维空间中第 i 个坐标轴的单位向量, 例如, $\boldsymbol{e}_1 = \{1, 0, \cdots, 0\}^{\mathrm{T}}$, $\boldsymbol{e}_n = \{0, 0, \cdots, 1\}^{\mathrm{T}}$. 像所有的极小化问题一样, 关键是找出从第 k 个迭代点 $\boldsymbol{x}_k$ 到下一个迭代点 $\boldsymbol{x}_{k+1}$ 的公式.

在探测移动中, 是用给定的步长向量 $\boldsymbol{\alpha}$ 用坐标轮换法进行搜索. 从 $\boldsymbol{x}_k$ 出发先沿第一个坐标轴方向取试验点 $\boldsymbol{x}_k^{(1)} = \boldsymbol{x}_k \pm \alpha_1 \boldsymbol{e}_1$, 由条件 $f(\boldsymbol{x}_k^{(1)}) < f(\boldsymbol{x}_k)$ 来决定上式右边的正负号; 若这样两次试验不能找到使函数值下降的 $\boldsymbol{x}_k^{(1)}$, 则令 $\boldsymbol{x}_k^{(1)} = \boldsymbol{x}_k$, 第一个坐标轴方向的探测移动完成. 再从 $\boldsymbol{x}_k^{(1)}$ 出发沿 $\boldsymbol{e}_2$ 方向作探测移动, 依次进行下去完成 n 个方向的探测移动后求得 $\boldsymbol{x}_k^{(n)}$. 若有 $\boldsymbol{x}_k^{(n)} \neq \boldsymbol{x}_k$, 则完成了这一次 (第 k 次) 迭代的探测移动; 反之, 若 $\boldsymbol{x}_k^{(n)} = \boldsymbol{x}_k$ 则没有找到目标函数值下降的方向, 应该缩小步长令 $\boldsymbol{\alpha} = \beta\boldsymbol{\alpha}, \beta$ 为小于 1 的正常数, 再进行第 k 次迭代的探测移动, 直到完成为止. 如果几次缩小步长, 使得 $\|\boldsymbol{\alpha}\| < \varepsilon$ 时仍满足 $\boldsymbol{x}_k^{(n)} = \boldsymbol{x}_k$, 则这时 $\|\nabla f(\boldsymbol{x}_k)\|$ 接近于 0, 迭代停止, $\boldsymbol{x}_k$ 取为近似极小点.

现假定第 k 次探测移动成功, 即 $\boldsymbol{x}_k^{(n)} \neq \boldsymbol{x}_k$, 那么 $\boldsymbol{x}_k^{(n)} - \boldsymbol{x}_k$ 可认为是目标函数在 $\boldsymbol{x}_k$ 点附近下降较快的方向, 下一步是进行模式移动. 从 $\boldsymbol{x}_k^{(n)}$ 点出发, 沿方向 $\boldsymbol{p}_k = \boldsymbol{x}_k^{(n)} - \boldsymbol{x}_k$ 以步长 $\lambda_k = 1$ 做一维搜索, 求出新的试验点 $\boldsymbol{x}'_{k+1}$:

$$\boldsymbol{x}'_{k+1} = \boldsymbol{x}_k^{(n)} + \lambda_k \boldsymbol{p}_k = 2\boldsymbol{x}_k^{(n)} - \boldsymbol{x}_k.$$

然后从 $\boldsymbol{x}'_{k+1}$ 出发进行探测移动, 得到点 $\boldsymbol{x}_{k+1}$. 这时有以下两种情况:

(1) $f(\boldsymbol{x}_{k+1}) < f(\boldsymbol{x}_k^{(n)})$, 则第 k 次迭代全部完成, 得到新迭代点 $\boldsymbol{x}_{k+1}$.

(2) $f(\boldsymbol{x}_{k+1}) \geqslant f(\boldsymbol{x}_k^{(n)})$, 说明模式移动得到的点不满足目标函数值下降这一要求, 因此, 重新从 $\boldsymbol{x}_k^{(n)}$ 点出发, 减小步长 $\boldsymbol{\alpha}$ 进行探测移动, 即在两次探测性移动之间没有作模式移动.

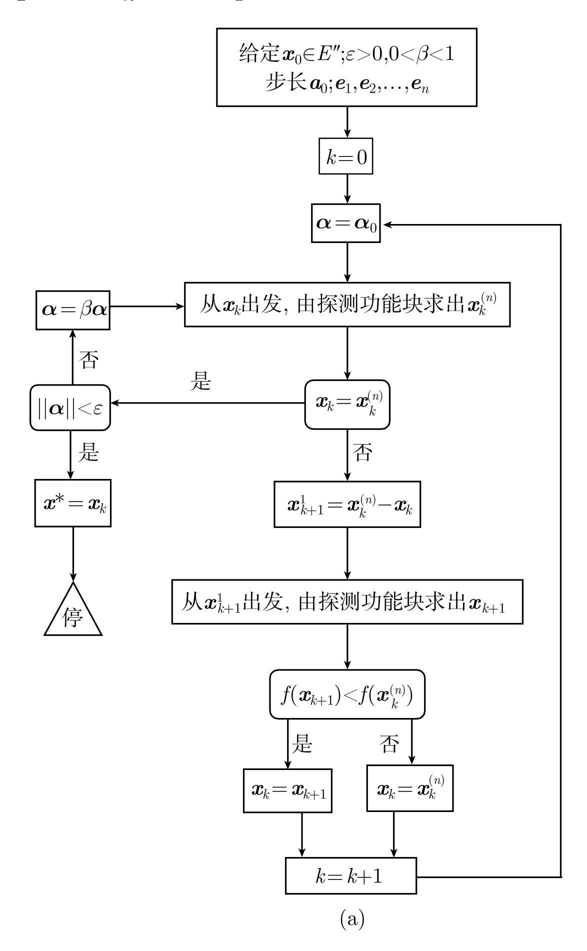
模式搜索法的程序框图如图 15.9 所示.

模式搜索法编制程序容易, 计算结果可靠, 但需要做大量的函数值计算. 它的收敛速度比较慢, 尤其在最优点附近是如此.

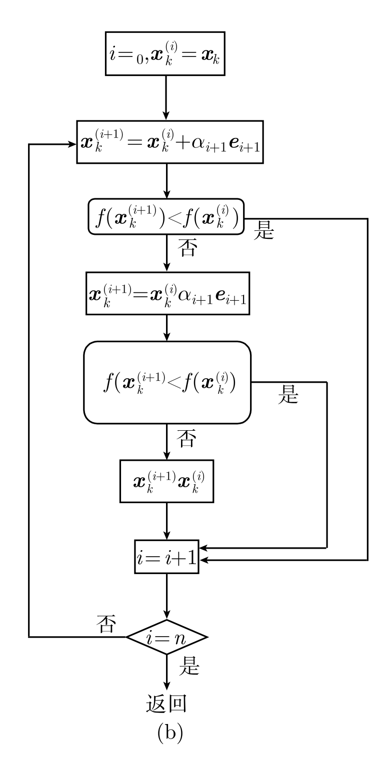
15.4.3　罗森布洛克转轴法

罗森布洛克 (Rosenbrock) 搜索方法的基本步骤如下：从某个初始点 $x_0 \in E^n$ 出发, 沿 n 维坐标单位向量 $\boldsymbol{e}_i(i=1,2,\cdots,n)$ 进行 n 次一维搜索, 得到下一个迭代点 $\boldsymbol{x}_1$, 这就是 15.4.1 节描述的坐标轮换法的一次迭代过程, 方向 $\boldsymbol{x}_1-\boldsymbol{x}_0$ 是目标函数下降较快的方向. 于是我们构造一组 n 个相互正交的单位向量 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 作为新的 "坐标" 方向, 其中 $\boldsymbol{p}_1=\boldsymbol{x}_1-\boldsymbol{x}_0$, 然后沿这组新的 "坐标" 方向再进行坐标轮换的 n 次搜索, 依次迭代下去, 直到满足某个迭代中止准则为止. 每进行一次迭代, 即每构造一组正交向量, 相当于在 n 维空间中将原来的坐标轴作转动, 所以这种方法又称为**转轴法**.

由以上叙述可见, 转轴法的关键在于找到某种方案, 根据第 k 个迭代点 $\boldsymbol{x}_k$, 以及由 $\boldsymbol{x}_{k-1}$ 求得的一组单位正交向量 $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\cdots$, $\boldsymbol{p}_n$ 来算出第 $\boldsymbol{k}$ 次迭代的一组正交单位向量 $\boldsymbol{p}_1'$, $\boldsymbol{p}_2'$, $\cdots$, $\boldsymbol{p}_n'$, 其中 $\boldsymbol{p}_1'$ 必定为



(a)



(b)

图 15.9　(a) 模式搜索法程序框图; (b) 探测功能块程序框图

$$\boldsymbol{p}_1' = \frac{\boldsymbol{x}_k - \boldsymbol{x}_{k-1}}{\|\boldsymbol{x}_k - \boldsymbol{x}_{k-1}\|}.$$

可以证明, 若 $\boldsymbol{p}_1'$, $\cdots$, $\boldsymbol{p}_n'$ 是线性无关的 n 个向量, $\underset{\sim}{A}$ 是 $n\times n$ 非奇异矩阵, 则运算

$$\boldsymbol{q}_j^{\mathrm{T}} = \boldsymbol{p}_j^{\mathrm{T}}\underset{\sim}{A}, \qquad j=1,\ 2,\ \cdots,\ n \tag{15.4.1}$$

得到的 n 个向量 $\boldsymbol{q}_1$, $\cdots$, $\boldsymbol{q}_n$ 一定是线性无关的. 而由下式表示的格拉姆–施密特 (Gram-Schmidt) 正交化过程, 可由线性无关的 n 个向量 $\boldsymbol{q}_1$, $\cdots$, $\boldsymbol{q}_n$ 产生一组 n 个正交单位向量 $\boldsymbol{p}_1'$, $\cdots$, $\boldsymbol{p}_n'$,

$$\begin{cases} \boldsymbol{p}_1^* = \boldsymbol{q}_1, \\ \boldsymbol{p}_j^* = \boldsymbol{q}_j + \displaystyle\sum_{i=1}^{j-1}\alpha_i\boldsymbol{p}_i^*, \qquad j=2,\ 3,\ \cdots,\ n, \\ \alpha_i = \dfrac{-\boldsymbol{q}_j^T\boldsymbol{p}_i}{(\boldsymbol{p}_i^*)\,\boldsymbol{p}_i}, \qquad i=1,\ 2,\ \cdots,\ j-1, \\ \boldsymbol{p}_j' = \boldsymbol{p}_j^*/\|\boldsymbol{p}_j^*\|, \qquad j=1,\ 2,\ \cdots,\ n. \end{cases} \tag{15.4.2}$$

现在设由 $\boldsymbol{x}_{k-1}$ 求得 $\boldsymbol{x}_k$ 的一组单位正交向量 (它们是线性无关的), 即式 (15.4.1) 中的 $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\cdots$, $\boldsymbol{p}_n$ 要使式 (15.4.2) 中的 $\boldsymbol{p}_1'$, $\cdots$, $\boldsymbol{p}_n'$ 能取作新的坐标方向, 则应有

$$\boldsymbol{p}_1^* = \boldsymbol{q}_1 = \boldsymbol{x}_k - \boldsymbol{x}_{k-1} \tag{15.4.3}$$

这可通过适当选取式 (15.4.1) 中的矩阵 $\underset{\sim}{A}$ 来达到

$$\underset{\sim}{A} = \begin{pmatrix} \lambda_1 & & & & \\ \lambda_2 & \lambda_2 & & 0 & \\ \lambda_3 & \lambda_3 & \lambda_3 & & \\ \vdots & \vdots & \vdots & & \\ \lambda_n & \lambda_n & \lambda_n & \cdots & \lambda_n \end{pmatrix}, \tag{15.4.4}$$

这里 λ_1, λ_2, $\cdots$, λ_n 是从 $\boldsymbol{x}_{k-1}$ 出发, 沿 $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\cdots$, $\boldsymbol{p}_n$ 做一维寻优达到 $\boldsymbol{x}_k$ 时的最优步长, 即

$$\boldsymbol{x}_k - \boldsymbol{x}_{k-1} = \lambda_1\boldsymbol{p}_1 + \lambda_2\boldsymbol{p}_2 + \cdots + \lambda_n\boldsymbol{p}_n.$$

若 $\boldsymbol{\lambda}_j \neq 0$, $j = 1, 2, \cdots, n$, 则明显有

$$\left|\underset{\sim}{A}\right| = \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n \neq 0,$$

即满足矩阵 $\underset{\sim}{\boldsymbol{A}}$ 为非奇异的要求. 此时, 根据式 (15.4.1), 有

$$\boldsymbol{p}_1^* = \sum_{i=1}^{n} \boldsymbol{p}_i A_{i1} = \lambda_1\boldsymbol{p}_1 + \cdots + \lambda_n\boldsymbol{p}_n = \boldsymbol{x}_k - \boldsymbol{x}_{k-1},$$

因此, 条件式 (15.4.3) 得到满足.

这样, 由前一次迭代中的正交单位矢量组 $\boldsymbol{p}_1$, $\cdots$, $\boldsymbol{p}_n$ 和从 $\boldsymbol{x}_{k-1}$ 到 $\boldsymbol{x}_k$ 的一维寻优决定的矩阵 $\underset{\sim}{A}$ 式 (15.4.4), 可通过式 (15.4.1), 式 (15.4.2) 求得下一次迭代中的正交单位向量组 $\boldsymbol{p}_1'$, $\cdots$, $\boldsymbol{p}_n'$.

沿一组正交向量 $\boldsymbol{p}_1$, $\cdots$, $\boldsymbol{p}_n$ 的一维寻优可按模式搜索法中 “探测移动” 的方式进行, 即若沿 $\boldsymbol{p}_i$ 方向按一定步长搜索的一次移动未能找到使函数值下降的点, 可缩小步长后再进行试验, 直到试验成功, 这时, $\lambda_i \neq 0$. 如果沿某个方向 $\boldsymbol{p}_j$ 不可能找到函数值下降的点, 则相应于 $\lambda_j = 0$, 这时算法便会中断, 必须进行修正. 设 n 个方向中共有 r 个方向 λ_i 等于 0, 则新的搜索方向 $\boldsymbol{p}_1'$, $\boldsymbol{p}_2'$, $\cdots$, $\boldsymbol{p}_n'$ 中保留这 r 个原来的单位向量 $\boldsymbol{p}_i(i = 1, 2, \cdots, r)$ 不变, 其余 $n-r$ 个向量 $\boldsymbol{p}_j(j = 1, \cdots, n-r)\lambda_j \neq 0$, 按上面描述的步骤正交化. 这 r 个原向量和 $n - r$ 个新向量之间彼此也是正交的, 可作为下一次迭代中的正交向量组.

如果在第 k 次迭代中, 用步长 $\boldsymbol{\alpha}$ 在 n 个正交方向做一维寻优, 步长 $\boldsymbol{\alpha}$ 已减小到 $\|\boldsymbol{\alpha}\| < \varepsilon$ 而各个方向都试验失败 (未能找到目标函数下降点), 则认为找到了极小值点的近似值 $\boldsymbol{x}_k \approx \boldsymbol{x}^*$.

转轴法的收敛速度比坐标轮换法有较大的改善, 特别是对于图 15.8(c) 所示存在脊线的目标函数, 它的收敛速度是比较快的, 解的稳定性也比较好; 但对高维目标函数, 其收敛速度下降.

15.4.4　单纯形法

单纯形法 1962 年由斯潘利 (Spendley), 赫克斯特 (Hext), 希姆斯沃思 (Himsworth) 等人提出, 1965 年内尔德 (Nelder), 米德 (Mead) 对它做了改进.

所谓**单纯形**是指 n 维空间中有 $n+1$ 个顶点的凸包, 例如, 一维空间中有两个顶点的直线段, 二维空间中有三个顶点的三角形, 三维空间中有四个顶点的四面体等 (如图 15.10).

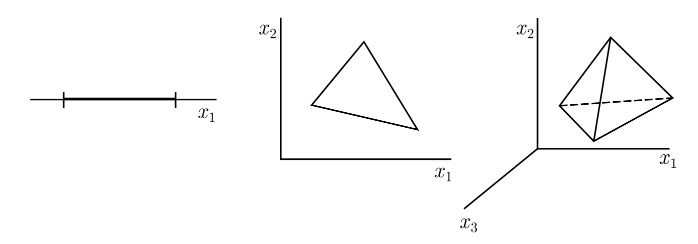


图 15.10　一、二、三维空间中的单纯形

单纯形法的基本思想是, 将 n 维空间中一个单纯形的 $n+1$ 个顶点处的目标函数值加以比较, 丢掉其中函数值最大的点, 并按一定规则进行探测性搜索, 寻找出一个函数值较低的新顶点, 与其余 n 个保留顶点构成新的单纯形. 如此迭代下去直到满足一定的迭代中止判据为止. 可见, 单纯形法包含两方面的问题, 首先是初始单纯形的构成, 其次是迭代过程.

首先讨论初始单纯形的构成. 以二维问题为例, 单纯形是三角形; 若三个顶点落在同一直线上, 则变成了一维问题, 即平面上找极小的问题变成在直线上找极小, 这种情形称为**退化**. 退化是我们所不希望的, 防止退化的方法是使三个顶点之间边长相等, 这样的三个顶点不可能落在同一直线上. 这一概念推广到 n 维空间的一般情形便得到**正规单纯形**的概念, 即任何两个顶点间距离都相等的 n+1 个点构成 n 维空间中的正规单纯形. 正规单纯形可以作为初始单纯形. 具体产生方法如下：对于给定初始点 $\boldsymbol{x}_0$ 和边长 a(两个顶点间的距离), 令

$$p = \frac{\sqrt{n+1}+n-1}{n\sqrt{2}}a, \quad q = \frac{\sqrt{n+1}-1}{n\sqrt{2}}a. \tag{15.4.5}$$

n 维正规单纯形 n+1 个顶点的坐标记为 $\boldsymbol{x}_0,\ \boldsymbol{x}_1,\ \cdots,\ \boldsymbol{x}_n$, 那么 $\boldsymbol{x}_j$ 的第 i 个分量 $x_j^{(i)}$ 可由下式求得：

$$\begin{cases} x_j^{(i)} = x_0^{(i)} + q, & i \neq j, \quad i,\ j = 1,\ 2,\ \cdots,\ n, \\ x_j^{(j)} = x_0^{(j)} + p, & j = 1,\ 2,\ \cdots,\ n. \end{cases} \tag{15.4.6}$$

下面讨论迭代过程, 即由已知的 $n+1$ 个顶点 $\boldsymbol{x}_0,\ \boldsymbol{x}_1,\ \cdots,\ \boldsymbol{x}_n$ 构成的单纯形, 比较各顶点的函数值, 去掉函数值最大的点, 按一定规则搜索求出函数值较低的新顶点, 与原来保留的 n 个顶点构成新的单纯形. 具体步骤如下：定义

$$\begin{aligned} &f(\boldsymbol{x}_h) = \max\{f(\boldsymbol{x}_0),\ f(\boldsymbol{x}_1),\ \cdots,\ f(\boldsymbol{x}_n)\}, \\ &f(\boldsymbol{x}_l) = \min\{f(\boldsymbol{x}_0),\ f(\boldsymbol{x}_1),\ \cdots,\ f(\boldsymbol{x}_n)\}, \\ &\boldsymbol{x}_m = \frac{1}{n}\left\{\sum_{i=0}^{n} \boldsymbol{x}_i - \boldsymbol{x}_h\right\}. \end{aligned}$$

$\boldsymbol{x}_m$ 的几何意义是除掉函数值最高的顶点 $\boldsymbol{x}_h$ 以外的其余 n 个顶点的位置重心. 迭代运算由下述四种运算组成：

(1) 反射. 令

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_m + \alpha(\boldsymbol{x}_m - \boldsymbol{x}_h),$$

$\boldsymbol{x}_{n+1}$ 是从 $\boldsymbol{x}_m$ 点出发沿 $\boldsymbol{x}_m - \boldsymbol{x}_h$ 方向 ($\boldsymbol{x}_h$ 指向 $\boldsymbol{x}_m$) 寻找单纯形的新顶点, $\alpha > 0$ 是给定的**反射系数**, $\boldsymbol{x}_{n+1}$ 是 $\boldsymbol{x}_h$ 关于 $\boldsymbol{x}_m$ 的反射点. 因为 $\boldsymbol{x}_h$ 是目标函数值最高的一个顶点, 所以 $\boldsymbol{x}_m - \boldsymbol{x}_h$ 方向是函数下降的方向, 沿该方向搜索可能找出函数值较低的点.

(2) 延伸. 如果反射成功, 即满足

$$f(\boldsymbol{x}_{n+1}) < f(\boldsymbol{x}_l),$$

说明沿 $\boldsymbol{x}_m - \boldsymbol{x}_h$ 方向搜索是正确的, 于是可以扩大搜索步长, 即将点再沿该方向作延伸, 取

$$\boldsymbol{x}_{n+2} = \boldsymbol{x}_m + \gamma(\boldsymbol{x}_{n+1} - \boldsymbol{x}_m),$$

其中 $\gamma > 1$ 为给定的**延伸系数**. 若延伸后有

$$f(\boldsymbol{x}_{n+2}) < f(\boldsymbol{x}_{n+1}),$$

则以点 $\boldsymbol{x}_{n+2}$ 代替原顶点 $\boldsymbol{x}_h$; 反之, 若上式不满足, 则以 $\boldsymbol{x}_{n+1}$ 代替原顶点 $\boldsymbol{x}_h$, 与保留的原来其余 n 个顶点构成新的单纯形.

(3) 收缩. 如果反射失败, 即 $f(\boldsymbol{x}_{n+1}) > f(\boldsymbol{x}_l)$, 则可分成几种情况：

(a) $f(\boldsymbol{x}_{n+1}) < f(\boldsymbol{x}_h')$

其中 $\boldsymbol{x}_h'$ 是 $\boldsymbol{x}_0$, $\boldsymbol{x}_1$, $\cdots$, $\boldsymbol{x}_n$ 中函数值仅次于 $f(\boldsymbol{x}_h)$ 的那个顶点. 这时以 $\boldsymbol{x}_{n+1}$ 代替原顶点 $\boldsymbol{x}_h$, 构成新的单纯形.

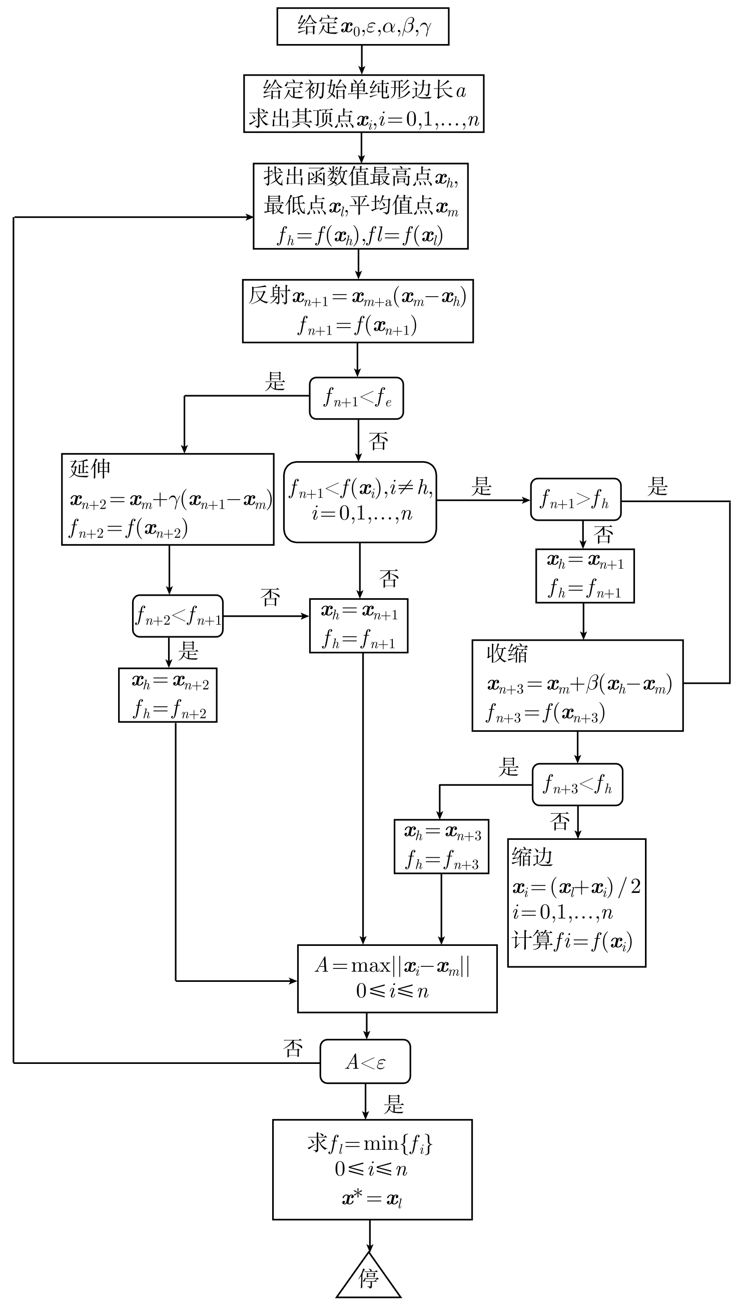


图 15.11　单纯形法程序框图

(b) $f(\boldsymbol{x}_{n+1}) \geqslant f(\boldsymbol{x}'_h)$

这时, 进行收缩运算. 又可分为两种情形:

$$\begin{cases} f(\boldsymbol{x}_{n+1}) \leqslant f(\boldsymbol{x}_h), \text{ 则 } \boldsymbol{x}_{n+3} = \boldsymbol{x}_m + \beta(\boldsymbol{x}_{n+1} - \boldsymbol{x}_m), \\ f(\boldsymbol{x}_{n+1}) > f(\boldsymbol{x}_h), \text{ 则 } \boldsymbol{x}_{n+3} = \boldsymbol{x}_m + \beta(\boldsymbol{x}_h - \boldsymbol{x}_m), \end{cases}$$

其中 $0 < \beta < 1$ 为给定的收缩系数. 上面两式的含义是在 $\boldsymbol{x}_{n+1}(\boldsymbol{x}_h)$ 和 $\boldsymbol{x}_m$ 的连线内搜索新的顶点以代替 $\boldsymbol{x}_h$, 如果 $f(\boldsymbol{x}_{n+3}) < f(\boldsymbol{x}_h)$, 表示收缩成功, 以 $\boldsymbol{x}_{n+3}$ 代替 $\boldsymbol{x}_h$ 构成新的单纯形.

(4) 缩边. 如果收缩失败, 即 $f(\boldsymbol{x}_{n+3}) \geqslant f(\boldsymbol{x}_h)$, 则应进行缩边, 保持点 $\boldsymbol{x}$ 的位置不动, 将所有向量 $\boldsymbol{x}_k - \boldsymbol{x}_l (k \neq l,\ k = 0,\ 1,\ \cdots,\ n)$ 的长度缩小一半, 用

$$\boldsymbol{x}_l + 0.5(\boldsymbol{x}_k - \boldsymbol{x}_l) = \frac{1}{2}(\boldsymbol{x}_l + \boldsymbol{x}_k), \qquad k \neq l,\ k = 0,\ 1,\ \cdots,\ n$$

代替原来的 $\boldsymbol{x}_k$, 然后再进行反射试验.

单纯形法的迭代终止判据可以是

$$|f(\boldsymbol{x}_h) - f(\boldsymbol{x}_l)| < \varepsilon$$

或

$$\max_{0 \leqslant k \leqslant n} \|\boldsymbol{x}_k - \boldsymbol{x}_m\| < \varepsilon,$$

这时以函数值最小的顶点 $\boldsymbol{x}_l$ 作为 $\boldsymbol{x}^*$ 的近似.

综合以上所述, 可给出图 15.11 所示的单纯形法程序框图.

单纯形法的优点是不必计算目标函数的梯度, 也不必按给定方向进行一维寻优, 每次迭代只需要计算一到二次函数值. 它的收敛速度依赖于参数 α, β, γ 的选择, 通常取 $\alpha = 1.0$, $\beta = 0.5$, $\gamma = 2$. 一般说来, 单纯形法是多维无约束极值问题的直接法中相当有效的一种.

15.5 最小二乘 Q^2 函数和似然函数的极值问题

以上介绍的极小化方法普遍适用于任意目标函数的极小化问题, 其中有些方法对目标函数的唯一要求是在极小值点附近能用二次函数来逼近. 为了推演更有效的极小化方法, 必须充分利用目标函数本身的特定性质. 在参数估计中我们看到, 最小二乘 Q^2 函数求极小和似然函数求极大是最小二乘法和极大似然法的核心. 如第 8 章和第 9 章所述, 似然方程和非线性模型的最小二乘 Q^2 函数极小化的方程往往没有解析解, 只能求助于本节所描述的数值迭代方法. 鉴于这两种参数估计方法在概率统计计算中的重要地位, 有必要利用 Q^2 函数和似然函数本身的性质, 导出更有效的求极值方法.

15.5.1 最小二乘 Q^2 函数极值

最小二乘法的核心问题是对 Q^2 函数求极小值 $Q^2_{\min}$. 在最通常遇到的最小二乘问题中, Q^2 函数一般可表示为函数平方和的形式

$$Q^2(\boldsymbol{x}) = \sum_{l=1}^{m} \varphi_l^2(\boldsymbol{x}). \tag{15.5.1}$$

例如, 式 (9.1.1), 式 (9.1.3)~式 (9.1.6) 都具有式 (15.5.1) 的形式, 其中 $\boldsymbol{x}$ 为待估计的参数. $Q^2(\boldsymbol{x})$ 等于极小值 $Q^2_{\min}$ 时的 $\boldsymbol{x}$ 值令为 $\boldsymbol{x}^*$, $\boldsymbol{x}^*$ 即为参数的最优解. 所以参数的最小二乘估计问题等价于目标函数

$$f(\boldsymbol{x}) \equiv Q^2(\boldsymbol{x}) = \sum_{l=1}^{m} \varphi_l^2(\boldsymbol{x}), \quad \boldsymbol{x} \in E^n,\ m \geqslant n \tag{15.5.2}$$

的极小化问题, 其中 m 为测量值个数, 它必须大于等于待估计参数 $\boldsymbol{x} = \{x_1, x_2, \cdots, x_n\}$ 的个数 n. Q^2 函数的极小化可利用牛顿法. 牛顿法的迭代公式中出现目标函数的二阶导数矩阵, 而对最小二乘问题, 由于目标函数具有平方和的形式, 可用 $\varphi_l(\boldsymbol{x})$ 的一阶导数来估计 $f(\boldsymbol{x})$ 的二阶导数, 这就使牛顿法具有更简单的形式, 称为**高斯–牛顿法**.

考察 $f(\boldsymbol{x})$ 的二阶导数矩阵 $\underset{\sim}{G}$, 它的 $n \times n$ 个元素为

$$\begin{aligned} G_{ij} &\equiv \frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\sum_{l=1}^{m}\varphi_l^2 = \frac{\partial}{\partial x_i}\sum_{l=1}^{m} 2\varphi_l \frac{\partial \varphi_l}{\partial x_j} \\ &= \sum_{l=1}^{m} 2\frac{\partial \varphi_l}{\partial x_i}\frac{\partial \varphi_l}{\partial x_j} + \sum_{l=1}^{m} 2\varphi_l \frac{\partial^2 \varphi_l}{\partial x_i \partial x_j} \quad i,\ j = 1,\ 2,\ \cdots,\ n. \end{aligned}$$

在所谓的**线性化近似**下, 上式右边含二阶导数的项比一阶导数项要小, 可以忽略, 这时, 有

$$G_{ij} \approx \sum_{l=1}^{m} 2\frac{\partial \varphi_l}{\partial x_i}\frac{\partial \varphi_l}{\partial x_j}, \qquad i,\ j = 1,\ 2,\ \cdots,\ n.$$

当矩阵 $\underset{\sim}{G}$ 是非奇异时, 它是正定矩阵.

记

$$g_{lj}(\boldsymbol{x}) = \frac{\partial \varphi_l}{\partial x_j}, \qquad l = 1,\ \cdots,\ m, \quad j = 1,\ \cdots,\ n, \tag{15.5.3}$$

即 $\underset{\sim}{\boldsymbol{g}}$ 为 $m \times n$ 矩阵, 则有

$$G_{ij} = 2\sum_{l=1}^{m} g_{li} g_{lj}, \qquad i,\ j = 1,\ 2,\ \cdots,\ n,$$

或写为

$$\underset{\sim}{G} = 2\,\underset{\sim}{\boldsymbol{g}}^{\mathrm{T}}\,\underset{\sim}{\boldsymbol{g}}\,. \tag{15.5.4}$$

又根据式 (15.5.2), 得

$$\frac{\partial f(\boldsymbol{x})}{\partial x_j} = 2\sum_{l=1}^{m}\varphi_l(\boldsymbol{x})\,\frac{\partial \varphi_l(\boldsymbol{x})}{\partial x_j}, \qquad j=1,\ 2,\ \cdots,\ n,$$

记

$$\boldsymbol{\phi}(\boldsymbol{x}) = \{\varphi_1(\boldsymbol{x}),\ \varphi_2(\boldsymbol{x}),\ \cdots,\ \varphi_m(\boldsymbol{x})\}^{\mathrm{T}}, \tag{15.5.5}$$

目标函数的梯度向量可表示为

$$\nabla f(\boldsymbol{x}) = 2\,\underset{\sim}{g}^{\mathrm{T}}\,\boldsymbol{\phi}(\boldsymbol{x}). \tag{15.5.6}$$

将式 (15.5.4), 式 (15.5.6) 代入牛顿法迭代公式 (15.3.17), 就得出高斯–牛顿法的迭代公式

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \left[\underset{\sim}{\boldsymbol{g}}_k^{\mathrm{T}}\,\underset{\sim}{\boldsymbol{g}}_k\right]^{-1}(\underset{\sim}{\boldsymbol{g}})_k^{\mathrm{T}}\,\boldsymbol{\phi}(\boldsymbol{x}_k). \tag{15.5.7}$$

高斯–牛顿法用一阶导数矩阵 $\underset{\sim}{\boldsymbol{g}}^{\mathrm{T}}\boldsymbol{\varphi}(\boldsymbol{x})$ 代替牛顿法中 $f(\boldsymbol{x})$ 的二阶导数矩阵, 从而大大减少了计算量. 但如同牛顿法一样, 对初始点 $\boldsymbol{x}_0$ 的选择要求比较严格, 要求与极小值点 $\boldsymbol{x}^*$ 相距不太远. 原因在于高斯–牛顿法利用了线性近似, 如与极小值点相距很远, 线性近似不再有效. 但应当指出, 如果 $\varphi_l(\boldsymbol{x})$ 是 $\boldsymbol{x}$ 的线性函数, 即 $f(\boldsymbol{x})$ 是 $\boldsymbol{x}$ 的二次函数 (这对应于最小二乘法的线性模型), 则二阶导数矩阵元 G_{ij} 中 $\dfrac{\partial^2\varphi_l}{\partial x_i\partial x_j}=0$, 这时线性化是严格正确的, 对初始迭代点 $\boldsymbol{x}_0$ 就没有特殊的要求.

与牛顿法的修正 (阻尼牛顿法) 相同, 在高斯–牛顿法中, 可视

$$\boldsymbol{p}_k = -\left[(\underset{\sim}{g})_k^{\mathrm{T}}(\underset{\sim}{g})_k\right]^{-1}(\underset{\sim}{g})_k^{\mathrm{T}}\,\boldsymbol{\phi}(\boldsymbol{x}_k) \tag{15.5.8}$$

为搜索方向, 利用阻尼牛顿法迭代公式

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k\boldsymbol{p}_k,$$

其中 λ_k 是使 $f(\boldsymbol{x}_k+\lambda\boldsymbol{p}_k)$ 达到极小的最优步长. 这一修正可进一步改善高斯–牛顿法的收敛速度.

15.5.2　似然函数极值

设总体 Y 的概率密度 $f(y|\boldsymbol{x})$ 函数形式为已知, $\boldsymbol{x}=\{x_1,\ x_2,\ \cdots,\ x_n\}$ 为待估计参数. $\boldsymbol{Y}=\{Y_1,\ Y_2,\ \cdots,\ Y_m\}$ 为总体 Y 的容量 m 的子样. 极大似然原理告诉我

们 (见第 8 章), 当似然函数

$$L(\boldsymbol{Y}|\boldsymbol{x}) = \prod_{l=1}^{m} f(Y_l|\boldsymbol{x})$$

或它的对数

$$\ln L(\boldsymbol{Y}|\boldsymbol{x}) = \sum_{l=1}^{m} \ln f(Y_l|\boldsymbol{x}) \equiv \sum_{l=1}^{m} \ln f_l(\boldsymbol{x})$$

取极大值时的参数值 $\boldsymbol{x}^*$ 即为待估计参数的最优估计. 因此, 极大似然法相当于使目标函数

$$F(\boldsymbol{x}) = -\sum_{l=1}^{m} \ln f_l(\boldsymbol{x}), \qquad \boldsymbol{x} \in E^n \tag{15.5.9}$$

极小化.

现在来考虑这种特定形式函数的二阶导数矩阵 $\underset{\sim}{G}$ 的 $n \times n$ 个元素, 显然,

$$\begin{aligned} G_{ij} =& \frac{\partial^2 F(\boldsymbol{x})}{\partial x_i \partial x_j} = -\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \sum_{l=1}^{m} \ln f_l(\boldsymbol{x}) \\ =& -\frac{\partial}{\partial x_i} \sum_{l=1}^{m} \frac{1}{f_l(\boldsymbol{x})} \cdot \frac{\partial f_l(\boldsymbol{x})}{\partial x_j} \\ =& \sum_{l=1}^{m} \frac{1}{f_l^2(\boldsymbol{x})} \cdot \frac{\partial f_l(\boldsymbol{x})}{\partial x_i} \cdot \frac{\partial f_l(\boldsymbol{x})}{\partial x_j} \\ & - \sum_{l=1}^{m} \frac{1}{f_l(\boldsymbol{x})} \cdot \frac{\partial^2 f_l(\boldsymbol{x})}{\partial x_i \partial x_j}. \end{aligned}$$

如同高斯–牛顿法中一样, 在函数 $f_l(\boldsymbol{x})$ 的线性近似下, 二阶导数项比一阶导数项小, 可予忽略, 因此, 有

$$G_{ij} \approx \sum_{l=1}^{m} \frac{1}{f_l^2(\boldsymbol{x})} \frac{\partial f_l(\boldsymbol{x})}{\partial x_i} \frac{\partial f_l(\boldsymbol{x})}{\partial x_j}, \qquad i,\ j = 1,\ 2,\ \cdots,\ n. \tag{15.5.10}$$

目标函数的一阶导数可由下法求得：显然,

$$\frac{\partial F(\boldsymbol{x})}{\partial x_j} = -\sum_{l=1}^{m} \frac{1}{f_l(\boldsymbol{x})} \frac{\partial f_l(\boldsymbol{x})}{\partial x_j}, \qquad j = 1,\ \cdots,\ n.$$

定义 $m \times n$ 矩阵 $\underset{\sim}{\boldsymbol{g}}$ 的元素为

$$g_{lj}(\boldsymbol{x}) = \frac{1}{f_l(\boldsymbol{x})} \frac{\partial f_l(\boldsymbol{x})}{\partial x_j}, \qquad j = 1,\ \cdots,\ n,\ l = 1,\ \cdots,\ m. \tag{15.5.11}$$

$$h_j(\boldsymbol{x}) = \sum_{l=1}^{m} g_{lj}(\boldsymbol{x}), \qquad j = 1, \ \cdots, \ n. \tag{15.5.12}$$

则目标函数的梯度向量可表示为

$$-\nabla F(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x}) \equiv [h_1(\boldsymbol{x}), \ h_2(\boldsymbol{x}), \ \cdots, \ h_n(\boldsymbol{x})]^{\mathrm{T}}, \tag{15.5.13}$$

而二阶导数矩阵 $\underset{\sim}{G}$ 可表示为

$$\underset{\sim}{G} = \underset{\sim}{\boldsymbol{g}}^{\mathrm{T}} \underset{\sim}{\boldsymbol{g}}. \tag{15.5.14}$$

若 $\underset{\sim}{G}$ 为非奇异, 则总是正定的.

将目标函数 $F(\boldsymbol{x})$ 的一阶导数和二阶导数的表达式 (15.5.13), 式 (15.5.14) 代入牛顿法的迭代公式 (15.3.17), 得到

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \left[\underset{\sim k}{\boldsymbol{g}}^{\mathrm{T}} \underset{\sim k}{\boldsymbol{g}}\right]^{-1} \boldsymbol{h}_k(\boldsymbol{x}). \tag{15.5.15}$$

因此, 在线性化近似下, 利用 $\underset{\sim k}{\boldsymbol{g}}^{\mathrm{T}} \cdot \underset{\sim k}{\boldsymbol{g}}$ (一阶导数运算) 代替二阶导数矩阵, 减少了计算量, 加快了迭代速度.

显然, 这里也可利用阻尼牛顿法来改善收敛速度, 其步骤与 15.5.1 节末所述的相类似.

15.6　局部极小和全域极小

迄今为止, 本章所介绍的极小化方法都是寻找局部极小值的, 至于所找到的极小是否为变量域的全域极小, 或者除此之外是否还存在其他局部极小, 上述方法不能给出任何信息. 如果目标函数有多个局部极小, 用前面介绍的方法求出的极小值点不能保证是最接近于初始点的那个局部极小值点. 事实上, 在使用这些方法时, 通常假定目标函数在待求极小值点的区间内是只含一个极小值点的单峰函数.

因此, 当目标函数在待求极小的区间内存在多个极小时, 就产生一个问题：我们要求的究竟是什么样的极小值点? 通常存在四种情形：

(1) 要求任意一个极小值的位置;

(2) 要求全域极小;

(3) 要求一个特定的极小值点, 它是待解问题的物理上正确的解, 但不一定是全域极小;

(4) 要求包含全域极小在内的所有局部极小.

可能性 (1) 是很少遇见的, 但最容易处理, 因为任何一种极小化算法都可以满足要求.

可能性 (2) 比较常见, 因为许多问题都希望求出 "费用" 最小的解, 这正是典型的全域极小值问题. 虽然有一些寻找全域极小的算法 (如文献 [121]、[122]), 但都没有指出终止迭代的适当判据. 对于变量空间存在上下界, 或者寻找一定区域内的全域极小这种特殊情况, 下面介绍的**网格法**和**随机搜索法**可以寻找全域极小, 它们都属于直接搜索方法.

在实际问题中, 可能性 (3) 最为常见. 这时极小值点 $\boldsymbol{x}^*$ 中某几个分量的近似值往往大体已经知道, 而所要寻找的解离这些近似值不远. 我们可以应用下述技巧来保证求出所需要的局部极小：将已知的近似分量值固定, 对其余分量进行极小化求出近似最优点, 然后从该点出发, 对所有分量进行极小化搜索.

可能性 (4) 这类问题是最困难的, 一种可能的方法是在 n 维空间中等间距地设置许多初始点, 从这些点出发逐一地搜索局部极小, 显然, 这是代价极高的做法.

15.6.1 网格法

设目标函数为 $f(\boldsymbol{x})$, $\boldsymbol{x} \in E^n$, 假定 $\boldsymbol{x}$ 的取值范围为已知

$$a_i \leqslant x_i \leqslant b_i, \qquad i = 1,\ 2,\ \cdots,\ n. \tag{15.6.1}$$

网格法就是在上式给定的 n 维区间内打网格, 求出所有网格点上目标函数的值并进行比较, 选出函数值最小的点. 然后在该点附近的区间加密网格 (减小网格点之间的距离) 再找出函数值最小的点, 如此循环下去, 直到相邻网格点之间的距离都小于控制误差 ε, 函数值最小的位置便是极小点的近似解. 网格可以是等间距的, 也可以是不等间距的, 以均匀网格较为简便. 为了不使全域极小被遗漏, 初始网格不能太稀, 网格间距应在计算中根据目标函数的行为加以调整. 显然, 对于高维目标函数, 网格点数目增长很快. 因而工作量极大, 实际上极少采用.

15.6.2 随机搜索法

最简单的随机搜索法称为**随机跳跃法**. 它的基本思想是在式 (15.6.1) 限定的变量 $\boldsymbol{x}$ 的区域内随机地均匀打点, 并比较这些随机点的函数值, 取其最小者为极小值点的近似解. 令 r 是 [0, 1] 区间内均匀分布的随机数列, 每产生一组 n 个随机数 $r_i (i = 1, 2, \cdots, n)$, 可得到符合式 (15.6.1) 条件的 n 维空间中的一个随机点

$$x_i = a_i + (b_i - a_i) r_i, \qquad i = 1,\ 2,\ \cdots,\ n. \tag{15.6.2}$$

共产生 J 组随机点, 并比较它们的目标函数值, 其最小者对应的随机点坐标可取作近似极小值点. 为了使近似极小值点有比较好的精度, J 值应当足够大, 这样便产生 nJ 个随机数. 显而易见, 这种方案对于高维目标函数是十分费时的.

采用**随机走步法**可以提高搜索效率. 从给定的初值 $\boldsymbol{x}_0$ 开始迭代过程. 由 $\boldsymbol{x}_k$ 产生 $\boldsymbol{x}_{k+1}$ 的公式是

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda \boldsymbol{U}_k, \tag{15.6.3}$$

其中 λ 是给定的初始搜索步长, $\boldsymbol{U}_k$ 是单位随机向量, 它定义为

$$\boldsymbol{U} = \frac{1}{\sqrt{\sum_{i=1}^{n} \eta_i^2}} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}, \qquad \|\boldsymbol{U}\| = 1, \tag{15.6.4}$$

其中 η_i 是 $[-1, +1]$ 区间内均匀分布随机数. 比较 $f(\boldsymbol{x}_{k+1})$ 和 $f(\boldsymbol{x}_k)$, 若 $f(\boldsymbol{x}_{k+1}) < f(\boldsymbol{x}_k)$, 则第 k 次迭代成功, 继续进行第 $k+1$ 次迭代; 反之, 若 $f(\boldsymbol{x}_{k+1}) \geqslant f(\boldsymbol{x}_k)$, 试验失败, 则随机地选择新的单位随机向量 $\boldsymbol{U}_k$, 直到第 $\boldsymbol{k}$ 次迭代成功为止. 如果经过多次随机向量的选择仍不能取得试验的成功, 则应当减小步长, 再进行试验.

如果到第 m 次迭代, 搜索步长已小于给定的控制常数 ε, 但经多次随机向量选择仍不能找到 $\boldsymbol{x}_{m+1}$ 满足

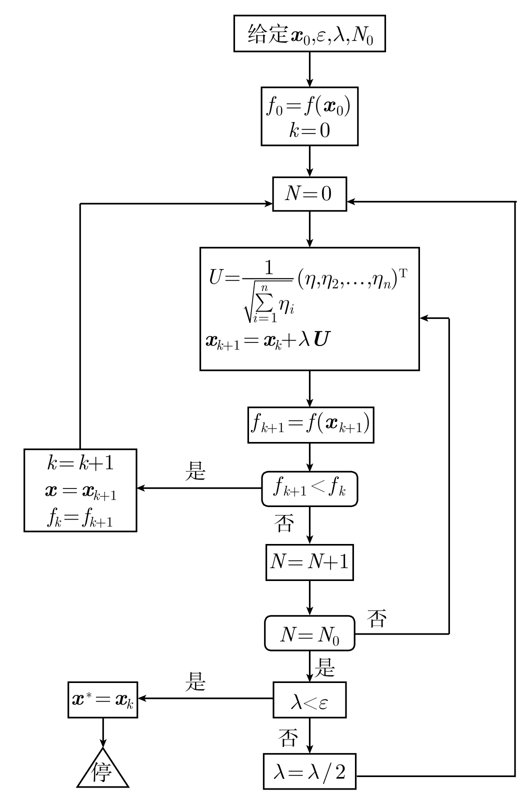


图 15.12　随机走步法程序框图

$$f(\boldsymbol{x}_{m+1}) < f(\boldsymbol{x}_m),$$

则 $\boldsymbol{x}_m$ 取为近似极小点.

由以上描述的步骤, 可给出图 15.12 所示随机走步法的程序框图.

上述随机走步法采用了固定步长 λ, 也可以在搜索过程中改变步长. 如果沿某个随机方向 $\boldsymbol{U}_k$ 的试验是成功的 (目标函数值下降), 可以沿该方向加大步长搜索, 或者沿该方向求最优步长, 即求满足下式的步长 λ_k:

$$f(\boldsymbol{x}_{k+1}) = f(\boldsymbol{x}_k + \lambda_k \boldsymbol{U}_k) = \min_{\lambda} f(\boldsymbol{x}_k + \lambda \boldsymbol{U}_k),$$

得到下一迭代点 $\boldsymbol{x}_{k+1}$, 这有助于加快向极小值点收敛的速度.

网格法和随机搜索法都是直接搜索法, 因而具有直接方法的优点和缺点, 两者相比较, 随机搜索法收敛速度比网格法快. 但对于高维目标函数, 即使用随机搜索法, 计算量仍然极大. 克服这一缺点的途径之一是: 首先用随机搜索法确定全域极小值点所在的大致范围, 然后在此范围内用收敛速度快的极小化方法求出全域极小的较精确的近似值.

15.7　约束 n 维极值问题

无约束极小化问题的解的容许区域是整个 n 维空间中的任意点, 在实际问题中, 往往对于解有一定的约束条件. 如本章引言所述, 约束 n 维极小化问题可表述为

$$\min f(\boldsymbol{x}) = f(\boldsymbol{x}^*), \qquad \boldsymbol{x} \in E^n, \tag{15.7.1}$$

并满足约束方程

$$h_i(\boldsymbol{x}) = 0, \qquad i = 1,\ 2,\ \cdots,\ l, \quad l < n, \tag{15.7.2}$$

$$g_j(\boldsymbol{x}) \geqslant 0, \qquad j = 1,\ 2,\ \cdots,\ m. \tag{15.7.3}$$

例如, 常常遇到下列几种对变量 $\boldsymbol{x}$ 的限制条件:

(1) 简单常数限

$$a_i \leqslant x_i \leqslant b_i, \qquad i = 1,\ 2,\ \cdots,\ n, \quad a_i,\ b_i\text{为常数}.$$

(2) 简单变量限

$$l_i(\boldsymbol{x}) \leqslant x_i \leqslant h_i\,(\boldsymbol{x}), \qquad i = 1,\ 2,\ \cdots,\ n.$$

(3) 隐含变量限

$$u_i(\boldsymbol{x}) \leqslant \nu_i(\boldsymbol{x}) \leqslant \omega_i\,(\boldsymbol{x}), \qquad i = 1,\ 2,\ \cdots,\ m.$$

情形 (1)、(2) 可视为 (3) 的特殊情形, 而 (3) 可改写为

$$\omega_i(\boldsymbol{x})-\nu_i(\boldsymbol{x})-u_i(\boldsymbol{x})\geqslant 0,\qquad \nu_i(\boldsymbol{x})-u_i(\boldsymbol{x})\geqslant 0,$$

$$i=1,\ 2,\ \cdots,\ m.$$

即具有式 (15.7.3) 的形式.

约束极小化问题的求解途径之一, 是对原目标函数 $f(\boldsymbol{x})$ 进行适当的修改, 得到新的目标函数 $F(\boldsymbol{x})$, 使得 $F(\boldsymbol{x})$ 的无约束极小解就是 $f(\boldsymbol{x})$ 的约束极小解. 显然, $F(\boldsymbol{x})$ 的构成与约束条件有关. 以下一些方法在求解约束极值问题中是常见的: ①利用等式约束方程消去一部分变量; ②引入拉格朗日乘子; ③通过变量代换消去约束; ④引入罚函数.

消去变量法①只适用于等式约束. 式 (15.7.2) 所示的 l 个等式约束使得 $\boldsymbol{x}=\{x_1,\ x_2,\ \cdots,\ x_n\}$ n 个变量中的 l 个可以用其余 $n-l$ 个变量来表示, 因此问题就化为 $n-l$ 个变量的无约束极小化问题. 方法②称为拉格朗日乘子法, 引入拉格朗日乘子 $\boldsymbol{\lambda}=\{\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_l\}$, 构造新的目标函数

$$F(\boldsymbol{x},\ \boldsymbol{\lambda})=f(\boldsymbol{x})+\sum_{i=1}^{l}\lambda_i h_i(\boldsymbol{x}),$$

于是式 (15.7.1)、式 (15.7.2) 的等式约束极小化问题化成 $F(\boldsymbol{x},\boldsymbol{\lambda})$ 的 $n+l$ 个变量无约束极小化问题. 方法①、②我们在 9.7 节和 9.8 节中讨论约束最小二乘估计问题时已做过介绍. 下面我们讨论方法③、④求解约束极值问题.

15.7.1 变量代换法

该方法的主要思想是将变量 $\boldsymbol{x}$ 变换成新变量 $\boldsymbol{y}$, 通过这一变换消除约束. 这样, 目标函数 $f(\boldsymbol{x})$ 的约束极小问题化为 $f(\boldsymbol{y})$ 的无约束极小问题, 变量 $\boldsymbol{y}$ 的极小值点 $\boldsymbol{y}^*$ 对应的变量值 $\boldsymbol{x}^*$ 即为 $f(\boldsymbol{x})$ 的约束极小值解. 一般, 变量代换法仅适用于一些简单的约束问题.

几种简单的约束条件适用的变量代换列举如下:

(1) 约束条件 $0\leqslant x_i\leqslant\infty,\ i=1,\ 2,\ \cdots,\ m$.

下列两种变量代换均适用:

$$x_i=y_i^2 \text{ 或 } x_i=\mathrm{e}^{y_i}.$$

(2) 约束条件 $0\leqslant x_i\leqslant 1,\ \ i=1,\ 2,\ \cdots,\ m$.

适用的变量代换是

$$x_i=\sin^2 y_i \text{ 或 } x_i=\frac{\mathrm{e}^{y_i}}{\mathrm{e}^{y_i}-\mathrm{e}^{-y_i}}$$

(3) 约束条件 $a_i \leqslant x_i \leqslant b_i$, $i = 1, 2, \cdots, m$ (a_i, b_i 常数).

适用的变量代换是

$$x_i = a_i + (b_i - a_i)\sin^2 y_i.$$

(4) 约束条件 $0 \leqslant x_i \leqslant x_j \leqslant x_k \leqslant \infty$.

适用的变换为

$$x_i = y_i^2, \qquad x_j = y_i^2 + y_j^2, \qquad x_k = y_i^2 + y_j^2 + y_k^2.$$

显然, 情况 (1)、(2)、(3) 中如果对搜索区间作相应的限制就相当于考虑了约束条件, 仍可用无约束极小化方法求解.

15.7.2 罚函数法

对于比较复杂的约束, 想通过变量代换来消去约束一般是做不到的. 而且在许多情形中, 拉格朗日函数关于 $\boldsymbol{x}$ 的极小不存在, 因此, 拉格朗日乘子法也不可行. 这时需用罚函数法求解. 它的基本思想是将目标函数 $f(\boldsymbol{x})$ 增广为一个新的函数 $F(\boldsymbol{x}, M)$, 称为**罚函数**

$$F(\boldsymbol{x}, M) = f(\boldsymbol{x}) + P(\boldsymbol{x}, M), \tag{15.7.4}$$

使约束极小化问题式 (15.7.1)~式 (15.7.3) 等价于罚函数 $F(\boldsymbol{x}, M)$ 的无约束极小化问题, 上式中 $P(\boldsymbol{x}, M)$ 称为罚项, M 称为罚因子.

约束条件式 (15.7.2)、式 (15.7.3) 相当于规定了问题**解的容许域**. 如果罚函数的选择使得极小值的搜索点列是从容许域外向极小值点逼近, 称为**外点法**; 反之, 搜索点列从容许域内向极小值点逼近, 则称为**内点法**.

首先介绍外点法, 它的罚函数形式取为

$$F(\boldsymbol{x}, M) = f(\boldsymbol{x}) + M\sum_{i=1}^{l}|h_i(\boldsymbol{x})|^{\alpha} + M\sum_{j=1}^{m}|\min[0, g_j(\boldsymbol{x})]|^{\beta}. \tag{15.7.5}$$

上式右边第二项当约束条件式 (15.7.2) 得到满足时等于 0, 否则恒大于 0, 它是该约束条件的罚项. 右边第三项中 $\min[0, g_i(\boldsymbol{x})]$ 表示在 $g_j(\boldsymbol{x})$ 和 0 之间选择小者, 即

$$\min[0, g_j(\boldsymbol{x})] = \begin{cases} 0, & 若 g_j(\boldsymbol{x}) \geqslant 0 (满足约束), \\ g_j(x), & 若 g_j(\boldsymbol{x}) < 0 (不满足约束). \end{cases}$$

因此, 第三项当不满足约束条件式 (15.7.3) 时恒大于 0, 为该约束的罚项. 罚因子 M 取为很大的正数, 指数 α, β 通常取为 2.

显然, 当不满足约束条件时, 罚项 $P(\boldsymbol{x}, M)$ 将很大; 仅当满足约束条件时, 罚函数 $F(\boldsymbol{x}, M)$ 才等于原目标函数 $f(\boldsymbol{x})$. 因此, 罚项的引入相当于对不遵守约束条件的

惩罚, 这样, $f(\boldsymbol{x})$ 在约束条件下的极小化问题就化为式 (15.7.5) 的罚函数 $F(\boldsymbol{x},M)$ 的无约束极小化问题. $F(\boldsymbol{x},M)$ 极小化的最优解 $\boldsymbol{x}^*(M)$ 与罚因子 M 有关, M 值越大, $\boldsymbol{x}^*(M)$ 越接近于 $f(\boldsymbol{x})$ 的约束极小化解. 也就是说, $M\to\infty$ 时, $\boldsymbol{x}^*(M)$ 收敛于 $f(\boldsymbol{x})$ 的约束极小化解. 在实际的迭代过程中, 第一步迭代选择一个罚因子 M 的值, 随着迭代次数的增多, 罚因子越取越大, 当 M 充分大时, 最优解 $\boldsymbol{x}^*(M)$ 趋近于约束极小值点.

如果对不同的约束需要加不同的权因子, 则可利用下式的罚函数:

$$F(\boldsymbol{x},\ \boldsymbol{M})=f(\boldsymbol{x})+\sum_{i=1}^{l}M_i\left|h_i(\boldsymbol{x})\right|^{\alpha}+\sum_{j=1}^{m}M_j\left|\min[0,\ g_j(\boldsymbol{x})]\right|^{\beta},\tag{15.7.6}$$

式中 $\boldsymbol{M}$ 为 $m+l$ 个分量的矢量, 每个分量都为正值, 它们的数值大小表示不同约束条件的权因子.

外点法的具体迭代步骤如下:

(1) 给定初始点 $\boldsymbol{x}_0$, 误差控制常数 $\varepsilon>0$; 常数 $C>1$, $M_0>0$(对于加权约束应给出 $M_0^{(i)}>0,\ i=1,\ 2,\ \cdots,\ l+m$); 令 $k=0$.

(2) 从 $\boldsymbol{x}_k$ 出发, 对式 (15.7.5) 或式 (15.7.6) 定义的函数 $F(\boldsymbol{x}_k,\boldsymbol{M}_k)$ 求出无约束极小值点, 作为下一迭代点 $\boldsymbol{x}_{k+1}$.

(3) 若罚项 $P(\boldsymbol{x}_{k+1},\boldsymbol{M}_k)<\varepsilon$, 则 $\boldsymbol{x}_{k+1}$ 为问题的近似解. 否则令 $M_{k+1}=CM_k$, 以 $k+1$ 代替 k, 返回步骤 (2).

由罚项的选择方法可知, 当 $\boldsymbol{x}$ 在解的容许域内时, $P(\boldsymbol{x},\boldsymbol{M})=0$; 当 $\boldsymbol{x}$ 在解的容许域外, 则有 $P(\boldsymbol{x},\boldsymbol{M})>0$. 在上述迭代过程中, 随着迭代次数 k 和罚因子 $\boldsymbol{M}_k$ 的增大, 罚项 $P(\boldsymbol{x}_k,\boldsymbol{M}_k)$ 由很大的正数逐渐趋近于 0, 因而点列 $\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots$ 是在解的容许域外部逼近问题最优解的, 这就是称为外点法的原因.

在外点法中, 当在某个充分大的 $\boldsymbol{M}_k$ 处终止迭代时, 近似最优点 $\boldsymbol{x}_k^*$ 一般只能近似地满足约束条件, 但 $\boldsymbol{x}_k^*$ 仍在解的容许域之外. 在某些实际问题中, 这样的近似解不能满足要求, 而希望近似解必须在解的容许域之内. 达到这一点的合理思路是构造适当的罚函数, 使得迭代点在约束条件限定的容许域内部逐渐逼近问题的最优解, 这就是**内点法**的基本思想.

设约束极小化问题为

$$\begin{gathered}\min f(\boldsymbol{x})=f(\boldsymbol{x}^*),\qquad \boldsymbol{x}\in E^n,\\ g_j(\boldsymbol{x})\geqslant 0,\qquad j=1,\ 2,\ \cdots,\ m.\end{gathered}\tag{15.7.7}$$

定义罚函数

$$F(\boldsymbol{x},\ r)=f(\boldsymbol{x})+r\sum_{j=1}^{m}\frac{1}{g_j(\boldsymbol{x})}\equiv f(\boldsymbol{x})+B(\boldsymbol{x},\ r).\tag{15.7.8}$$

其中 $r>0$ 是罚因子, $B(\boldsymbol{x},r)$ 是罚项. 如对不同的约束需加不同的权因子, 则定义罚函数为

$$F(\boldsymbol{x},\ \boldsymbol{r})=f(\boldsymbol{x})+\sum_{j=1}^{m}\frac{r_j}{g_j(\boldsymbol{x})}\equiv f(\boldsymbol{x})+B(\boldsymbol{x},\ \boldsymbol{r}). \tag{15.7.9}$$

由以上三式可知, 在约束条件限定的解的容许域内, 罚项恒大于 0; 当罚项趋近于 0 时, $F(\boldsymbol{x},\boldsymbol{r})$ 趋近于原目标函数 $f(\boldsymbol{x})$. 同时, 在求解 $F(\boldsymbol{x},\boldsymbol{r})$ 的无约束极小时, 若迭代点列 $\{\boldsymbol{x}_k\}$ 从容许域内趋近约束边界, $g_j(\boldsymbol{x}_k)$ 趋近于 0, 罚项迅速增大, 这就有效地阻止了 $\boldsymbol{x}_k$ 越出解的容许域. 因此, 式 (15.7.7) 所示的约束极小化问题就化为 $F(\boldsymbol{x},\boldsymbol{r})$ 的无约束极小化问题, $\boldsymbol{r}\to 0$ 时, $F(\boldsymbol{x},\boldsymbol{r})$ 的无约束最优解就是约束极小化问题的解. 并且由以上所述可知, 内点法中迭代的初始点 $\boldsymbol{x}_0$ 必须选在解的容许域内, 迭代点列 $\{\boldsymbol{x}_k\}$ 是在容许域内收敛向问题的最优解的.

内点法的具体迭代步骤如下:

(1) 给定初始点 $\boldsymbol{x}_0$, 满足 $g_j(\boldsymbol{x}_0)>0\{j=1,2,\cdots,m\}$; 控制常数 $\varepsilon>0$, 常数 $\boldsymbol{r}_0>0(0<C<1)$; 令 $k=0$.

(2) 以 $\boldsymbol{x}_k$ 为初始点, 求 (15.7.8) 或 (15.7.9) 中罚函数 $F(\boldsymbol{x}_k,\boldsymbol{r}_k)$ 的无约束极小, 作为下一迭代点 $\boldsymbol{x}_{k+1}$.

(3) 若 $B(\boldsymbol{x}_{k+1},\ \boldsymbol{r}_k)<\varepsilon$, 则 $\boldsymbol{x}_{k+1}$ 为问题 (15.7.7) 的解, 终止迭代; 否则令 $r_{k+1}=r_k/C$, 以 $k+1$ 代替 k, 返回步骤 (2).

内点法中罚函数的罚项也可取为约束函数的对数, 即罚函数的形式是

$$F(\boldsymbol{x},\ \boldsymbol{r})=f(\boldsymbol{x})-r\sum_{j=1}^{m}\ln g_j(\boldsymbol{x}),$$

$$F(\boldsymbol{x},\ \boldsymbol{r})=f(\boldsymbol{x})-\sum_{j=1}^{m}r_j\ln g_j(\boldsymbol{x}), \tag{15.7.10}$$

这称为**对数罚函数法**. 当 $g_j(\boldsymbol{x})\to 0_+$, 项 $-\ln g_j(\boldsymbol{x})\to\infty$, 罚项变得很大, 在阻挡迭代点列 $\{\boldsymbol{x}_k\}$ 接近和越过约束方程限定的容许域边界这一方面, $-\ln g_j(\boldsymbol{x})$ 的作用与 $[g_j(\boldsymbol{x})]^{-1}$ 是相似的.

内点法的优点在于每次迭代点都是容许点, 当迭代到一定次数求出问题的近似解, 一定落在解的容许域内. 但内点法要求初始迭代点位于容许域内部, 有时做到这一点是困难的; 同时内点法不能处理等式约束. 为了弥补这些不足, 可以将内点法和外点法结合起来使用. 当初始点 $\boldsymbol{x}_0$ 给定之后, 对等式约束和不被 $\boldsymbol{x}_0$ 满足的那些不等式约束用外点法罚项 $P(\boldsymbol{x},M)$, 而对被 $\boldsymbol{x}_0$ 所满足的那些不等式约束则用内点法罚项 $B(\boldsymbol{x},r)$, 即罚函数定义为

$$F(\boldsymbol{x},\ r)=f(\boldsymbol{x})+\sum_{i=1}^{l}|h_i(\boldsymbol{x})|^{\alpha}+\frac{1}{r}\sum_{j\in J_1}|\min[0,\ g_j(x)]|^{\beta}-r\sum_{j\in J_2}\ln g_j(x) \tag{15.7.11}$$

其中下标集合 J_1, J_2 定义为

$$J_1 = \{j | g_j(\boldsymbol{x}_0) \leqslant 0, \quad i = 1, \cdots, m\},$$
$$J_2 = \{j | g_j(\boldsymbol{x}_0) > 0, \quad i = 1, \cdots, m\},$$

$\sum\limits_{j \in J_1}$ 表示对所有属于 J_1 的那些项求和.

这种方法称为**混合罚函数法**, 其迭代过程与内点法相同, 而初始点 $\boldsymbol{x}_0$ 不必满足 $g_j(\boldsymbol{x}_0) > 0$ 的条件.

罚函数法适用范围较广, 对于复杂的等式、不等式约束都能较好地处理, 使用也很方便. 它的缺点是必须求一系列的无约束极小, 计算量很大; 同时参数 M_0、r_0、C 的选取对收敛速度有较大的影响.

例 15.1 共振产生的确定

粒子反应

$$\pi^- \mathrm{p} \to \pi^- \mathrm{p} \pi^+ \pi^-$$

中实验测定了 $\pi^+\pi^-$ 系统的不变质量谱 (不变质量的定义见例 3.2), 测定表明, 在 765MeV 和 1260MeV 处事例数出现两个峰, 这表明产生了粒子共振 $\rho(765)$ 和 $f(1260)$. 实验测定的质量谱可用下述理论谱描写：

$$q(M, \boldsymbol{\alpha}) = \alpha_\rho BW_\rho(M) + \alpha_f BW_f(M) + \alpha_b B(M), \tag{15.7.12}$$

其中 M 是 $\pi^+\pi^-$ 系统的不变质量, $BW_\rho(M)$、$BW_f(M)$ 是描写 ρ 共振和 f 共振的布雷特–维格纳 (Breit-Wigner) 函数, $B(M)$ 是本底项, 这三个函数都是归一化并且已知的. ρ 共振、f 共振和本底项的相对比率 α_ρ、α_f 和 α_b 是待定的未知参数.

由问题的性质可知, 参数 α_ρ、α_f 和 α_b 必须满足以下约束：

$$0 \leqslant \alpha_\rho, \ \alpha_f, \ \alpha_b \leqslant 1, \qquad \alpha_\rho + \alpha_f + \alpha_b = 1. \tag{15.7.13}$$

其中的等式约束是相对比率的归一化要求. 由该等式约束可消去一个待估计参数, 即将三个参数中的一个用其余两个参数来表示, 假定利用 $\alpha_b = 1 - \alpha_\rho - \alpha_f$ 消去 α_b, 余下的 α_ρ、α_f 为待估计参数. 设用最小二乘法做参数估计, 令实验测定值与式 (15.7.12) 的理论值之间的离差平方和函数 (Q^2 函数) 记为 $f(\alpha_\rho, \alpha_f)$, 则问题变为求满足约束条件

$$0 \leqslant \alpha_\rho, \ \alpha_f \leqslant 1, \qquad 0 \leqslant 1 - \alpha_\rho - \alpha_f \leqslant 1 \tag{15.7.14}$$

的函数 $f(\alpha_\rho, \alpha_f)$ 的极小值点 α_ρ^*、α_f^*.

利用罚函数内点法求解, 为此必须将上式的约束条件改写为式 (15.7.7) 的形式

$$\begin{cases} \alpha_\rho \geqslant 0, \\ 1-\alpha_\rho \geqslant 0, \\ \alpha_f \geqslant 0, \\ 1-\alpha_f \geqslant 0, \\ 1-\alpha_\rho-\alpha_f \geqslant 0. \end{cases} \tag{15.7.15}$$

式 (15.7.14) 中的约束条件 $1 \geqslant 1-\alpha_\rho-\alpha_f$ 等价于 $\alpha_\rho+\alpha_f \geqslant 0$, 这一要求已由 $\alpha_\rho \geqslant 0$ 和 $\alpha_f \geqslant 0$ 自动满足了. 根据式 (15.7.8), 罚函数形式应为

$$F(\boldsymbol{\alpha},\ r)=f(\alpha_\rho,\ \alpha_f)+r\left(\frac{1}{\alpha_\rho}+\frac{1}{1-\alpha_\rho}+\frac{1}{\alpha_f}+\frac{1}{1-\alpha_f}+\frac{1}{1-\alpha_\rho-\alpha_f}\right). \tag{15.7.16}$$

利用无约束极小化方法求得极小值点近似解 α_ρ^*、α_f^*, 并由式 (15.7.13) 求得 $\alpha_b=1-\alpha_\rho^*-\alpha_f^*$. 由此得出了 ρ 共振、f 共振和本底项三者的相对比率, 问题得解.

15.8　参数的误差估计

就参数估计而言, 我们对于极小化问题的兴趣不是函数极小值本身, 而在于求得使目标函数 $f(\boldsymbol{x})$ 极小化的一组变量 (参数值)$x_1^*,\ x_2^*,\ \cdots,\ x_n^*$, 这组值就是待估计参数的最优值. 在求得了最优值之后, 有待解决的问题是怎样计算这些参数值的不确定性 (误差).

如果参数是用极大似然法估计的, 根据 8.4.2 节和 8.4.3 节的讨论, 当子样容量 n 很大, 似然函数为渐近的多维正态分布或极大似然估计量是充分、有效估计量时, 参数 $\boldsymbol{\vartheta}$ 的极大似然估计 $\hat{\boldsymbol{\vartheta}}$ 的方差矩阵有简单的形式 (式 (8.4.7) 和式 (8.4.12))

$$V_{ij}^{-1}(\boldsymbol{\vartheta})=-\frac{\partial^2 \ln L}{\partial\vartheta_i\partial\vartheta_j}\bigg|_{\vartheta=\hat{\vartheta}},$$

如果是最小二乘估计问题, Q^2 是参数 $\boldsymbol{\vartheta}$ 的二次函数, 则有 (式 (9.9.11))

$$V_{ij}^{-1}(\hat{\boldsymbol{\vartheta}})=\frac{1}{2}\left(\frac{\partial^2 Q^2}{\partial\vartheta_i\partial\vartheta_j}\right)_{\vartheta=\hat{\vartheta}},$$

而且该式对非线性模型、Q^2 不是 $\boldsymbol{\vartheta}$ 的二次函数的情形式也近似适用. 当使用极小化方法的记号, $\boldsymbol{\vartheta}$ 对应于变量 $\boldsymbol{x}$, $\hat{\boldsymbol{\vartheta}}$ 对应于 $\boldsymbol{x}^*$, 目标函数 $f(\boldsymbol{x})$ 对应于 $-\ln L$(极大似然法) 和 $Q^2(\boldsymbol{x})$(最小二乘法), 因此有

$$V_{ij}^{-1}(\boldsymbol{x}^*) \propto \frac{\partial^2 f}{\partial x_i\partial x_j}\bigg|_{x=x^*}. \tag{15.8.1}$$

可见 $\boldsymbol{x}^*$ 的协方差矩阵的逆阵正好是目标函数的二阶导数矩阵在极小值点 $\boldsymbol{x}^*$ 处的值. $\boldsymbol{x}^*$ 的各分量 $x_1^*, x_2^*, \cdots, x_n^*$ 的误差由 $\boldsymbol{x}^*$ 协方差矩阵各对角元素的平方根表示, 这样, 问题就归结为求黑塞矩阵的逆在 $\boldsymbol{x}^*$ 处的值 $\underset{\sim}{H}^{-1}(\boldsymbol{x}^*)$.

对于在算法中已给出黑塞矩阵的极小化方法, 如牛顿法、共轭方向法, 参数误差的确定就十分简单. 在变尺度法中, 当目标函数为二次函数, 方向矩阵 $\underset{\sim}{A}(\boldsymbol{x}^*)$ 收敛于 $\underset{\sim}{H}^{-1}(\boldsymbol{x}^*)$; 对于一般的非二次目标函数, 可在每个迭代点用二次函数作为目标函数的近似, 因而 $\underset{\sim}{A}(\boldsymbol{x}^*)$ 也收敛于 $\underset{\sim}{H}(\boldsymbol{x}^*)$ 的逆阵. 这样, 参数的误差可由 $\underset{\sim}{A}(\boldsymbol{x}^*)$ 的对角元素求得.

对于无法求出黑塞矩阵的极小化方法, 参数误差的确定要寻求别的途径. 记得在极大似然法和最小二乘法中, 可利用图像法来确定参数的误差. 即超表面 $\ln L(\boldsymbol{\vartheta})$ 与超平面 $\ln L = \ln L_{\max} - a$(极大似然法) 或超表面 $Q^2(\boldsymbol{\vartheta})$ 与超平面 $Q^2 = Q^2_{\min} + a$(最小二乘法) 相截所围成的区域对应于一定误差的置信域 (8.6 节、9.9 节). 例如, 对于一个参数的特定情况, 利用极小化方法的记号, 最小二乘法中一个标准差的置信区间 $[x_1, x_2]$ 可由下式确定 ($f(x) = Q^2$):

$$f(x_1) = f(x_2) = f(x^*) + 1;$$

类似地, 对于极大似然法, 则是

$$f(x_1) = f(x_2) = f(x^*) + 0.5$$

($f(x) = -\ln L(x)$). 这种方法可以推广, 用以对一般目标函数的参数最优值作误差估计, 即找出满足

$$f(x_1) = f(x_2) = f(x^*) + a \tag{15.8.2}$$

的点 x_1、x_2, 其中常数 a 的选择是使得能给出所需要的置信区间 (见 8.6 和 9.9 节的讨论).

在一般情形下, 目标函数 $f(x)$ 在极小值 x^* 附近可用二次函数 $\varphi(x) = A + Bx + Cx^2$ 作为近似, 这时满足式 (15.8.2) 的 x_1、x_2 是一元二次方程的两个根, 因而容易求出. 如果目标函数比较复杂, 在极小值点附近不能用二次函数作为近似, 则 x_1、x_2 的确定就不那么简单.

在西欧核子中心 (CERN) 的计算机程序库极小化程序包 MINUIT 中[130], 参数的误差是用下述技巧寻找的：设已求出目标函数 $f(\boldsymbol{x})$ 的极小值点近似值 $\boldsymbol{x}^* = \{x_1^*, x_2^*, \cdots, x_n^*\}$, 对应的函数值 $f(\boldsymbol{x}^*)$ 记为 f^0, 则 x_i^* 的上端误差 $\sigma_{\rm h}(x_i^*)$ 和下端误差 $\sigma_{\rm l}(x_i^*)$ 由下式确定：

$$f(x_1^*, \cdots, x_{i-1}^*, x_i^* + \sigma_{\rm h}(x_i^*), x_{i+1}^*, \cdots, x_n^*) = f^0 + a,$$

$$f(x_1^*,\ \cdots,\ x_{i-1}^*,\ x_i^*-\sigma_1(x_i^*),\ x_{i+1}^*,\ \cdots,\ x_n^*)=f^0+a,$$
$$i=1,\ 2,\ \cdots,\ n. \tag{15.8.3}$$

考察图 15.13, 图中函数极小值点为 x_i^* (只考虑 x_i 的变化, 其余分量皆为极小值点保持不变). 首先给定一个上端误差的初值 $\sigma_h^0(x_i^*)$, 由此确定一条二次曲线 (抛物线)

$$f(x_i)=f^0+2(\sigma_h^0)^{-2}\cdot(x_i-x_i^*)^2,$$

它与直线 $f=f^0+a$ 相交于点 B, 由 B 点引一条垂直于 x_i 轴的直线与目标函数曲线 $f(x_i)$ 相交于 $B\prime$ 点. 然后可构成一条经过极小值 A 以及 B' 的抛物线, 它与直线 $f=f^0+a$ 相交于 C, 与 C 具有相同 x_i 值的函数点为 C'. 再构成一条通过 $AB'C'$ 的抛物线, 与直线 $f=f^0+a$ 交于 D, 相应的函数点为 D'. 然后由 $B'C'D'$ 构成抛物线, 与直线 $f=f^0+a$ 交于 E, 相应的函数点为 E'. 如此迭代下去, 由最后三个函数点求出一个新函数点, 直到求出一个新的函数点 K', 满足

$$\left|f(x'_K)-(f^0+a)\right|<\varepsilon$$

为止, ε 是某个给定的常数, 这时从 A 到 K 之间的距离即为所求的 $\sigma_h(x_i^*)$. 通过类似的步骤向 x_i 减小的方向可求得 $\sigma_1(x_i^*)$. 这种方法可以估计目标函数行为相当奇特时的参数误差, 但显然计算过程是比较费时的.

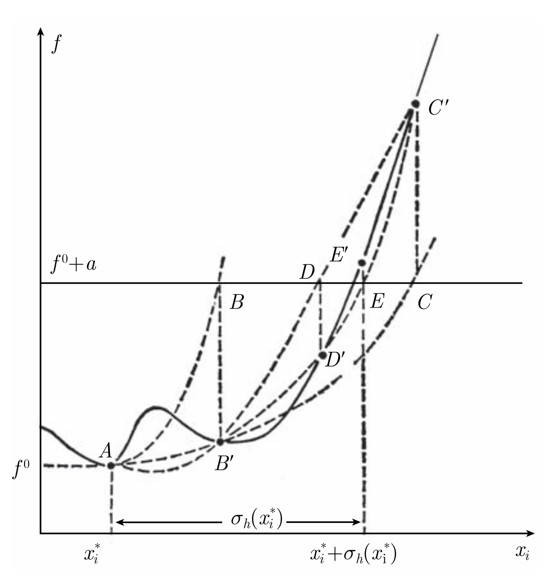


图 15.13　参数误差的确定

第 16 章　去弥散方法

在 4.20 节中我们已经提到, 对于所研究的随机变量, 利用某种仪器或装置进行测量, 由于实验测量存在探测效率 (一般小于等于 1) 和测量误差, 故实验测量得到的数据 (实验分布) 不能直接反映随机变量原有的分布 (原分布). 一般而言, 原分布是所研究的物理量所服从的物理规律的反映, 因而求得原分布是实验研究希望达到的目标. 校正实验测量导致的畸变, 从实验分布导出原分布的过程称为**去弥散** (unfolding), 数学上也称之为**去卷积** (deconvolution) 或**反演问题** (inverse problems). 去弥散方法在粒子物理和核物理、粒子和核天体物理、宇宙线物理和宇宙学、光学和医学成像重建、射电天文学和结晶学等领域都得到了广泛的发展和应用.

本章中讨论的去弥散方法本质上属于经典统计范畴, 许多内容取自于 [137], [138], [139], [160], [161] 等文献. 但其中某些方法也属于贝叶斯统计, 例如 16.5.3 节介绍的基于贝叶斯统计的极大熵原理, 参见文献 [85], [154], [155], [145].

我们在 16.1 节中, 讨论去弥散问题的数学表述以及符号的定义. 16.2 节讨论利用响应矩阵的求逆方法去弥散. 这一方法在实际中很少使用, 但它是别的更好的求解方法的一个出发点. 16.3 节讨论的修正因子法是一种比较简单的、易于实行的去弥散方法, 但其效果依赖于所研究过程和探测器的蒙特卡罗模拟. 本章的主要内容, 即正规化去弥散方法 (regularized unfolding) 将在 16.4 节 ~16.7 节加以阐述, 包括寻解的策略, 若干正规函数的讨论, 解的方差和偏差 (bias) 的估计方法等内容. 在 16.8 节中讨论去弥散的若干例子. 16.9 节则讨论去弥散方法的数值计算.

应当强调指出, 许多问题中并不需要对测量得到的分布进行去弥散处理. 例如, 当我们的目的是为了检验测量结果与怎样的理论预期相一致, 我们可以假设理论预期具有某种分布, 将探测器造成的畸变效应包含进来, 就可以直接与测量值进行比较. 这一过程比对测量值进行去弥散处理后再与原理论分布比较要简单得多.

但是, 如果不做去弥散处理, 测量值就无法与其他实验的结果直接进行比较, 因为不同实验中实验分辨的效应一般是不同的. 还可能发生这样的情况：在实验测量完成许多年之后, 一种新的理论出现了, 而那时, 每个实验的响应函数或响应矩阵 (参见下文) 可能已经无法得知, 也即无法将实验分辨的效应加入到新的理论分布中去以与原来的实验分布进行比较了. 因此, 如果一个特别重要的实验测量结果需要保留下来, 则应当同时保留实验测量值及其响应矩阵. 可惜的是, 这种做法往往是不实际的, 也很少有人这样做.

通过对测量值分布的去弥散处理, 所得到的结果 (亦即原分布的一种估计) 能够与其他实验结果以及理论预期直接比较, 这是去弥散处理的必要性. 需要进行去弥散处理的原因还在于, 如图像重建这类的应用中, 在未经去弥散修正的实验分布中原图像的某些特征可能无法辨认; 只有经过去弥散处理后才能将图像的这些特征显现出来. 这对于需要用重建后的图像来确定下一阶段处理的工作 (例如医疗诊断), 显然是极其重要的.

本章中我们假定已经认定有必要进行去弥散处理, 待解决的问题是给出去弥散的方法并具体加以实现.

16.1　去弥散问题的数学表述

考虑一个随机变量 x, 我们希望确定它的概率密度函数. 本章中我们假定对于 x 的测量有有限的分辨, 其探测效率小于 100%, 并且存在本底过程. 作为一个例子, 我们可以考虑测量对象是放射性原子核 β 衰变发射的电子的能量分布, 即随机变量 x 相应于发射电子的能量.

所谓 “有限的分辨” 指的是, 由于测量误差的存在, x 的测量值会偏离其真值, 其偏离量为一随机量. 例如, 一次特定的 β 衰变发射的电子能量是某个确定值 (真值), 但由于测量装置存在分辨, 所记录到的电子能量一般与其真值不同. 因此, 每个观测事例要用两个量来表征：真值 y(未知量) 和观测值 x.

一般情形下我们还必须考虑到原分布的某些真值虽然确实存在, 但却可能没有其对应的测量值. 比如在 β 衰变的例子中, 由于探测器没有全部包围放射源周围的全部立体角, 因而衰变发射的某个电子会逃逸出探测器未被探测到, 因此该电子的能量就没有对应的测量值; 或者由于电子能量低于某个最低的阈值, 因而不能产生大到足以被探测到的信号, 也没有对应的测量值. 原分布变量真值 y 能够产生某个测量值的概率称为探测效率 $\varepsilon(y)$, 从上面的叙述可以看到, 它一般依赖于事例的真值 y.

设**真值** (原) **分布**的概率密度为 $f_{\text{true}}(y)$. 为了构建一个关于 $f_{\text{true}}(y)$ 的估计量, 必须利用一组参数来表示它. 如果对于 $f_{\text{true}}(y)$ 的函数形式没有先验知识, 那么可以用一个 M 个子区间的归一化直方图来表示 y 的分布. 子区间 j 中出现 y 的概率 p_j 等于该子区间内概率密度 $f_{\text{true}}(y)$ 的积分

$$p_j = \int_{\Delta y_j} f_{\text{true}}(y)\,\mathrm{d}y. \tag{16.1.1}$$

其中 Δy_j 代表子区间 j.

设我们进行一个实验, 其中原分布产生的总事例数为 m_{tot}. 一般而言, 观测到的事例数将与 m_{tot} 不同. 总事例数 m_{tot} 可以是一个确定值, 也可以是一个随机变

量. 这两种情形下, 事例总数的期望值我们都表示为 $\mu_{\rm tot} = E(m_{\rm tot})$, 于是子区间 j 中事例数的期望值为

$$\mu_j = \mu_{\rm tot} p_j. \tag{16.1.2}$$

我们用向量 $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_M)$ 来表示 "**真值直方图**". 应当指出, $\mu_i\,(i=1,\cdots,M)$ 并不是直方图各子区间中的实际事例数, 而是其对应的期望值, 即 μ_i 一般并不是整数. 例如, 可以将子区间 i 中的实际事例数视为一个均值 μ_i 的随机变量 m_i. 但是, 由于探测器的有限分辨和效率, m_i 并不是一个直接观测量, m_i 甚至不出现于去弥散问题的数学表述中. 我们将构建的估计量是直接针对参数 μ_i 的.

为了方便起见, 通常还需要构建观测值的直方图. 设我们有了 x 的测量值的一个样本, 构成一个 N 个子区间的直方图, 得到这 N 个子区间中的事例数 $\boldsymbol{n} = (n_1, \cdots, n_N)$, 称之为 "**观测直方图**".

一般地, 观测直方图的子区间数 N 可以大于、小于或等于真值直方图的子区间数 M. 设子区间 i 包含 n_i 个事例, 观测总事例数为 $\sum_i n_i = n_{\rm tot}$. 需要考虑两种情况. 第一种情形, 变量 n_i 视为期望值 ν_i 的相互独立的泊松变量. 这一模型下子区间 i 中观测到 n_i 个事例的概率为

$$P(n_i; \nu_i) = \frac{\nu_i^{n_i} {\rm e}^{-\nu_i}}{n_i!}. \tag{16.1.3}$$

因为若干个泊松变量之和仍是一个泊松变量 (见 4.3 节), 故 $n_{\rm tot}$ 服从期望值 $\nu_{\rm tot} = \sum_i \nu_i$ 的泊松分布. 第二种情形 $n_{\rm tot}$ 为一固定值, 这时 n_i 服从多项分布. 在这两种情形中, 期望值都用下式表示:

$$\nu_i = E(n_i). \tag{16.1.4}$$

基于极大似然法的去弥散方法中需要用到似然函数, 为此需要知道数据 $\boldsymbol{n} = (n_1, \cdots, n_N)$ 的概率分布的形式 (泊松分布、多项分布等) 来构建似然函数. 另一种选择是基于最小二乘法的去弥散方法, 这时需要知道数据 $\boldsymbol{n} = (n_1, \cdots, n_N)$ 的协方差矩阵

$$V_{ij} = {\rm cov}(n_i, n_j). \tag{16.1.5}$$

在本章中我们假定, 数据的概率分布的形式或者协方差矩阵为已知.

期望值 $\nu_i = E(n_i)$ 可表示为

$$\nu_i = \mu_{\rm tot} \int_{\Delta x_j} {\rm d}x \int {\rm d}y\, [s(x|y)\, \varepsilon(y)\, f_{\rm true}(y)]. \tag{16.1.6}$$

式中 $s(x|y)$ 是给定真值为 y 并且事例被观测到的条件下, 测量值为 x 的条件概率, Δx_i 表示子区间 i, $s(x|y)$ 是归一化函数即 $\int s(x|y)\,{\rm d}x = 1$. 我们称 s 为**分辨函数**,

在图像重建的应用中则称为**点弥散函数** (point spread function). **响应函数**定义为：

$$r(x|y) = s(x|y)\,\varepsilon(y), \tag{16.1.7}$$

它表示给定真值 y 并且探测效率 $\varepsilon \in [0,1]$ 条件下, 观测值为 x 的概率. 需要指出, 响应函数 $r(x|y)$ 不是归一化函数. 在实际的测量中, 真值分布 f_{true} 被响应函数所**弥散**, 因此从实验测量分布来估计 f_{true} 的过程被称为**去弥散**.

将式 (16.1.6) 对于 y 的积分拆分成所有子区间中的求和运算, 并将分子和分母同时乘以 μ_j, 子区间 i 中观测事例数的期望值变成

$$\begin{aligned}\nu_i &= \sum_{j=1}^{M} \frac{\displaystyle\int_{\Delta x_i} \mathrm{d}x \int_{\Delta y_j} \mathrm{d}y \left[s(x|y)\,\varepsilon(y)\,f_{\text{true}}(y)\right]}{(\mu_j/\mu_{\text{tot}})}\mu_j \\ &= \sum_{j=1}^{M} R_{ij}\mu_j,\end{aligned} \tag{16.1.8}$$

其中**响应矩阵**$\underset{\sim}{R}$ 的元素由下式给定：

$$\begin{aligned}R_{ij} &= \frac{\displaystyle\int_{\Delta x_i} \mathrm{d}x \int_{\Delta y_j} \mathrm{d}y \left[s(x|y)\,\varepsilon(y)\,f_{\text{true}}(y)\right]}{\displaystyle\int_{\Delta y_j} \mathrm{d}y \left[f_{\text{true}}(y)\right]} \\ &= \frac{P\left(\text{真值在 } j \text{ 子区间, 观测值在 } i \text{ 子区间}\right)}{P\left(\text{真值在 } j \text{ 子区间}\right)} \\ &= P\left(\text{观测值在 } i \text{ 子区间} \mid \text{真值在 } j \text{ 子区间}\right).\end{aligned} \tag{16.1.9}$$

式中字母 P 表示概率, 因此响应矩阵元 R_{ij} 是真值 y 出现在 j 子区间条件下, 观测值 x 落在 i 子区间的条件概率. R 的非对角元素的效应是造成真值分布精细结构的模糊化. 真值直方图中主要集中于一个子区间的一个峰在实际观测中会分散到若干个子区间中. 真值直方图中彼此相隔若干个子区间的两个峰在一定条件下有可能在实际观测中合并为一个宽峰.

我们从式 (16.1.9) 的第一行可以看到, 响应矩阵依赖于概率密度 $f_{\text{true}}(y)$. 但它是先验地未知的, 因为我们的目的就是确定 $f_{\text{true}}(y)$. 如果真值直方图的子区间宽度足够小, 以至于 $s(x|y)$ 和 $\varepsilon(y)$ 在子区间宽度范围内可近似视为常数, 那么分子和分母对于 $f_{\text{true}}(y)$ 的直接依赖相互抵消. 在下文中我们将假定这种近似是成立的, 并且可以忽略由于 $f_{\text{true}}(y)$ 的不确定性导致的响应矩阵的误差.

虽然 $s(x|y)$ 和 $\varepsilon(y)$ 独立于出现特定 y 值的概率, 即独立于 $f_{\text{true}}(y)$, 但它们一般并不是完全模型无关的. 变量 y 不一定是能影响获得一个测量值 x 的概率的

唯一因素. 就 β 衰变的例子而言, y 代表电子的真实能量, 而 x 代表其测量能量. 由于探测器的不同部分可能有不同的分辨函数和探测效率, $s(x|y)$ 和 $\varepsilon(y)$ 一般将依赖于电子的角分布, 而 β 衰变的不同理论模型会对其角分布做出不同的预期.

在下文中我们将忽略这种模型依赖性, 直接假定分辨函数 $s(x|y)$ 和效率 $\varepsilon(y)$, 从而响应矩阵 R_{ij} 仅仅依赖于测量装置. 即我们假定, $\underset{\sim}{R}$ 能够几乎无误差地加以确定. 在实际情况中, $\underset{\sim}{R}$ 或者是利用事先已经知道真值 $\underset{\sim}{y}$ 的刻度实验加以确定, 或者是利用蒙特卡罗模拟确定, 后者依赖于对探测器中发生的物理过程的知识的精确程度. 实际问题中, 模型依赖可能不可忽略, 而且对于探测器的了解不可能尽善尽美, 这两者都需要考虑为系统误差的可能来源.

应当指出, 响应矩阵 R_{ij} 一般不是对称矩阵和方矩阵, 其中第一下标 $i=1,\cdots,N$ 标记观测直方图的子区间号, 第二下标 $j=1,\cdots,M$ 标记真值直方图的子区间号. 对第一个下标求和并利用关系式 $\int s(x|y)\mathrm{d}x=1$ 可得

$$\begin{aligned}\sum_{i=1}^{N}R_{ij}&=\sum_{i=1}^{N}\frac{\displaystyle\int_{\Delta x_i}\mathrm{d}x\int_{\Delta y_j}\mathrm{d}y s(x|y)\varepsilon(y)f_{\mathrm{true}}(y)}{(\mu_j/\mu_{\mathrm{tot}})}\\&=\frac{\displaystyle\int_{\Delta y_j}\mathrm{d}y\varepsilon(y)f_{\mathrm{true}}(y)}{\displaystyle\int_{\Delta y_j}f_{\mathrm{true}}(y)\mathrm{d}y}\\&\equiv\varepsilon_j\end{aligned}\tag{16.1.10}$$

即子区间 j 中效率的平均值.

除了有限的分辨和效率之外, 还必须容许测量装置在没有出现所研究的真实事例的情形下却存在测量值的可能性, 即测量值由某种**本底**过程所产生. 在 β 衰变的情形下, 这可能是由探测器中的杂散信号导致的结果, 例如样品中存在所研究的放射性核之外的放射性核, 又如存在宇宙线粒子导致的相互作用, 等等. 假设产生于本底过程的、子区间 i 中观测到的事例数期望值为 β_i, 则关系式 (16.1.8) 修改为

$$\nu_i=\sum_{j=1}^{M}R_{ij}\mu_j+\beta_i.\tag{16.1.11}$$

应当指出, β_i 包括了探测器的有限分辨和效率对于本底过程的效应. 它们一般根据刻度实验或者根据本底过程和探测器响应的蒙特卡罗模拟加以确定. 在以下的讨论中, 我们假定 β_i 值是已知值. 当然实际情形中 β_i 值只能确定到一定的精度, 因此本底的不确定性也成为去弥散结果的系统误差的来源之一.

小结一下, 我们有下述向量 (一般情况下也表示直方图或分布):

(1) 真值直方图 (每个子区间中原分布真实事例数的期望值) $\boldsymbol{\mu}=(\mu_1,\cdots,\mu_M)$;

(2) 归一化真值直方图 (概率) $\boldsymbol{p}=(p_1,\cdots,p_M)=\boldsymbol{\mu}/\mu_{\text{tot}}$;

(3) 观测事例数 (实验数据) $\boldsymbol{n}=(n_1,\cdots,n_N)$;

(4) 观测事例数期望值 $\boldsymbol{\nu}=(\nu_1,\cdots,\nu_N)$;

(5) 效率 $\boldsymbol{\varepsilon}=(\varepsilon_1,\cdots,\varepsilon_N)$;

(6) 本底期望值 $\boldsymbol{\beta}=(\beta_1,\cdots,\beta_N)$.

已经假定, 我们或者已经知道实验数据 $\boldsymbol{n}$ 的概率分布的形式, 能够用它来构建似然函数; 或者已经知道协方差矩阵 $V_{ij}=\text{cov}\,(n_i,n_j)$, 可以用它来构建 χ^2 函数. 此外, 我们还有了响应矩阵 R_{ij}, 其中 $i=1,\cdots,N$ 表示观测直方图的子区间, 而 $j=1,\cdots,M$ 表示真值直方图的子区间. 我们还假定 $\underset{\sim}{R}$ 和 $\boldsymbol{\beta}$ 皆为已知. 向量 $\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\beta}$ 和矩阵 $\underset{\sim}{R}$ 之间存在下述关系

$$\boldsymbol{\nu}=\underset{\sim}{R}\,\boldsymbol{\mu}+\boldsymbol{\beta} \tag{16.1.12}$$

这一矩阵方程中 $\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\beta}$ 应当理解为列向量. 我们的目的是构建对于真值直方图 $\boldsymbol{\mu}$ 的估计量 $\hat{\boldsymbol{\mu}}$, 或者构建对于概率 $\boldsymbol{p}$ 的估计量 $\hat{\boldsymbol{p}}$.

16.2　响应矩阵求逆法

本节中我们将讨论一种构造关于真值直方图向量 $\boldsymbol{\mu}$ 的估计量的方法. 这种方法表面上看起来是显而易见的, 但是, 它给出的解通常是不可取的. 考察真值和观测直方图子区间相同的情形, 即 $M=N$. 我们暂时假定矩阵方程 $\boldsymbol{\nu}=\underset{\sim}{R}\,\boldsymbol{\mu}+\boldsymbol{\beta}$ 可以求逆, 得到

$$\boldsymbol{\mu}=\underset{\sim}{R}^{-1}(\boldsymbol{\nu}-\boldsymbol{\beta})\,. \tag{16.2.1}$$

$\boldsymbol{\nu}$ 的估计量的一种显而易见的选择是由相对应的观测值给定：

$$\hat{\boldsymbol{\nu}}=\boldsymbol{n}. \tag{16.2.2}$$

于是立即可得 $\boldsymbol{\mu}$ 的估计量

$$\hat{\boldsymbol{\mu}}=\underset{\sim}{R}^{-1}(\boldsymbol{n}-\boldsymbol{\beta})\,. \tag{16.2.3}$$

容易证明这实际上就是对数似然函数求极大得到的解：

$$\ln L(\boldsymbol{\mu})=\sum_{i=1}^{N}\ln P(n_i;\nu_i), \tag{16.2.4}$$

式中 $P(n_i;\nu_i)$ 是泊松 (当事例总数 n_{tot} 为期望值 ν_{tot} 的泊松变量) 或二项 (事例总数 n_{tot} 为常数) 分布. 它也是使

$$\chi^2(\boldsymbol{\mu}) = \sum_{i,j=1}^{N} (\nu_i - n_i)\left(V^{-1}\right)_{ij}(\nu_j - n_j) \tag{16.2.5}$$

达到极小的最小二乘解. 由于关系式 $\boldsymbol{\nu} = \underset{\sim}{R}\boldsymbol{\mu} + \boldsymbol{\beta}$ 成立, $\ln L(\boldsymbol{\mu})$ 和 $\chi^2(\boldsymbol{\mu})$ 可以视为 $\boldsymbol{\mu}$ 或 $\boldsymbol{\nu}$ 的函数. 也就是说, 将式 (16.2.4) 或式 (16.2.5) 对 μ_i 求导数时, 用到关系式 $\partial\nu_i/\partial\mu_j = R_{ij}$ (见式 (16.1.11)).

我们来研究响应函数求逆得到的式 (16.2.3) $\boldsymbol{\mu}$ 的估计量的偏差和方差. $\hat{\mu}_j$ 的期望值由下式给定:

$$\begin{aligned} E(\hat{\mu}_j) &= \sum_{i=1}^{N}\left(\underset{\sim}{R}^{-1}\right)_{ji} E(n_i - \beta_i) = \sum_{i=1}^{N}\left(\underset{\sim}{R}^{-1}\right)_{ji}(\nu_i - \beta_i) \\ &= \mu_j, \end{aligned} \tag{16.2.6}$$

所以估计量 $\hat{\mu}_j$ 是无偏的, 因为按照假定 $\hat{\nu}_i = n_i$ 是无偏的. 对于协方差矩阵, 我们有

$$\begin{aligned} \operatorname{cov}(\hat{\mu}_i, \hat{\mu}_j) &= \sum_{k,l=1}^{N}\left(\underset{\sim}{R}^{-1}\right)_{ik}\left(\underset{\sim}{R}^{-1}\right)_{jl}\operatorname{cov}(n_k, n_l) \\ &= \sum_{k=1}^{N}\left(\underset{\sim}{R}^{-1}\right)_{ik}\left(\underset{\sim}{R}^{-1}\right)_{jk}\nu_k, \end{aligned} \tag{16.2.7}$$

其中, 我们用到了独立泊松变量的协方差矩阵 $\operatorname{cov}(n_k, n_l) = \delta_{kl}\nu_k$.

下面, 我们用记号 $V_{ij} = \operatorname{cov}(n_i, n_j)$ 表示数据的协方差矩阵, 记号 $U_{ij} = \operatorname{cov}(\hat{\mu}_i, \hat{\mu}_j)$ 表示真值分布估计量的协方差矩阵. 于是方程式 (16.2.7) 可写成矩阵形式:

$$\underset{\sim}{U} = \underset{\sim}{R}^{-1}\underset{\sim}{V}\left(\underset{\sim}{R}^{-1}\right)^T. \tag{16.2.8}$$

现在我们来说明为什么用响应矩阵求逆方法构建的估计量一般是不可取的, 或者说是不适合实际使用的. 考察图 16.1 所示的示例. 原真值分布 $\boldsymbol{\mu}$ 如图 16.1(a) 所示, 观测期望值 $\boldsymbol{\nu}$ 的分布则如图 16.1(b) 的直方图所示.

图 16.1 中, 直方图 $\boldsymbol{\nu}$ 按关系式 $\boldsymbol{\nu} = \underset{\sim}{R}\boldsymbol{\mu}$ 计算, 即本底 β 取为 0. 响应矩阵 $\underset{\sim}{R}$ 是按照标准偏差等于子区间宽度的 1.5 倍的高斯分辨函数和效率 ε_i 取为 1 计算得到的. 其结果是, 真值直方图中某子区间中的一个事例, 观测直方图中该子区间观测到一个事例的概率为 26%, 紧挨着的左 (右) 边的子区间内观测到一个事例的概率为 21%, 左 (右) 边第二个或更远的子区间内观测到一个事例的概率为 16%.

图 16.1(c) 显示的是数据直方图 $\boldsymbol{n} = (n_1, \cdots, n_N)$. 它们是利用图 16.1(b) 显示的 ν_i 作为均值的泊松分布通过蒙特卡罗方法产生的. 因为每一子区间中的事

例数的范围在 10^2 到 10^3 之间, 故 n_i 的相对统计误差 (标准差与均值之比) 在 3%~10%之间.

图 16.1(d) 显示的是按照方程式 (16.2.3) 对矩阵求逆得到的估计值 $\hat{\boldsymbol{\mu}}$. 误差杆标记每个子区间中的标准偏差. 与 n_i 的相对统计误差在 3%~10%之间极不相同, 各子区间中的 μ_j 起伏很大, 而且误差与估计值本身大小差不多 (注意, 该图纵坐标的尺度增大了). 相邻子区间估计量 $\hat{\mu}_j$ 的关联系数 $\mathrm{cov}(\hat{\mu}_i, \hat{\mu}_j)/\sigma_{\hat{\mu}_i}\sigma_{\hat{\mu}_j}$ 接近 -1.

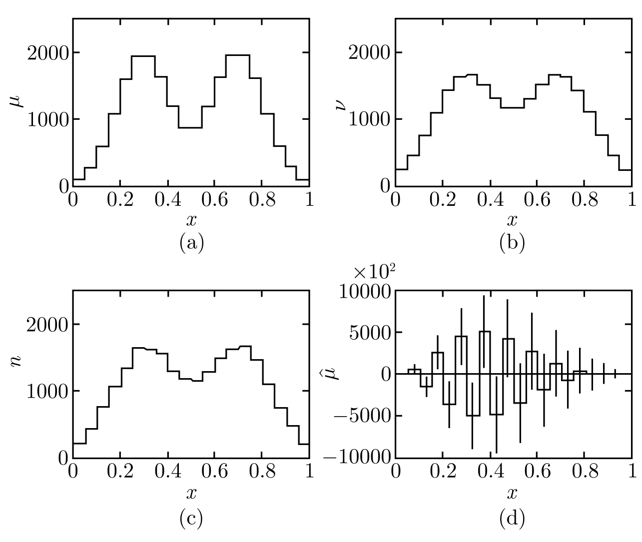


图 16.1　(a) 假设的真值直方图 $\boldsymbol{\mu}$; (b) 期望值直方图 $\boldsymbol{\nu} = \underset{\sim}{R}\boldsymbol{\mu}$; (c) 观测数据直方图 $\boldsymbol{n}$; (d) 由响应矩阵求逆得到的估计量 $\hat{\boldsymbol{\mu}}$

产生这种灾难性的失败的原因在于我们不知道期望值 ν; 而如果我们知道的话, 就可以直接计算 $\hat{\boldsymbol{\mu}} = \underset{\sim}{R}^{-1}\boldsymbol{\nu}$. 实际情况是, 我们只知道数据 $\boldsymbol{n}$, 并用它作为 $\boldsymbol{\nu}$ 的估计量 $\hat{\boldsymbol{\nu}}$, 但 $\boldsymbol{n}$ 是个随机变量因而具有统计涨落. 记得响应矩阵的效应是使得任何精细结构发生弥散. 如果在 $\boldsymbol{\mu}$ 中存在相距很近的峰, 尽管在 $\boldsymbol{\nu}$ 中它们会合并在一起, 但仍然能残留某些精细结构的 "遗迹". 当将 $\underset{\sim}{R}^{-1}$ 运算施加于 $\boldsymbol{\nu}$, 这些原始结构的遗迹有可能被恢复回来. 但各子区间的观测数据 $\boldsymbol{n} = (n_1, \cdots, n_n)$ 存在各自的统计涨落, 当用 $\boldsymbol{n}$ 作为 $\boldsymbol{\nu}$ 的估计量计算 $\boldsymbol{\mu}$ 的估计量 $\hat{\boldsymbol{\mu}}$, 会导致图 16.1(d) 中所示的大量 "虚假" 的精细结构.

可以将方程式 (16.2.8) 的协方差矩阵 $\underset{\sim}{U}$ 与 7.4 节的克拉美–罗不等式求得的最小方差界进行比较. 为此, 我们假定 n_i 为均值 ν_i 的相互独立的泊松变量. 于是似然函数为

$$\ln L(\mu) = \sum_{i=1}^{N} \ln P(n_i; \nu_i) = \sum_{i=1}^{N} \ln\left(\frac{\nu_i^{n_i} \mathrm{e}^{-\nu_i}}{n_i!}\right). \tag{16.2.9}$$

略去与 $\boldsymbol{\mu}$ 无关的项, 得到

$$\ln L(\boldsymbol{\mu}) = \sum_{i=1}^{N} (n_i \ln \nu_i - \nu_i). \tag{16.2.10}$$

令 $\ln L(\boldsymbol{\mu})$ 对于 $\boldsymbol{\mu}$ 的各分量的导数为 0, 并利用式 (16.1.11), 有

$$\frac{\partial \ln L}{\partial \mu_k} = \sum_{i=1}^{N} \left(n_i \frac{\partial \ln \nu_i}{\partial \nu_i} \frac{\partial \nu_i}{\partial \mu_k} - \frac{\partial \nu_i}{\partial \mu_k} \right) = \sum_{i=1}^{N} \left(\frac{n_i}{\nu_i} - 1 \right) R_{ik} = 0, \tag{16.2.11}$$

我们就得到前面已经用到过的估计量 $\hat{\boldsymbol{\nu}} = \boldsymbol{n}$. 再进一步求二阶导数得到

$$\frac{\partial^2 \ln L}{\partial \mu_k \partial \mu_l} = -\sum_{i=1}^{N} \frac{n_i R_{ik} R_{il}}{\nu_i^2}. \tag{16.2.12}$$

对于无偏估计量情形下协方差矩阵逆阵 (见式 (8.3.19)) 的最小方差界于是为

$$\begin{aligned} \left(\underset{\sim}{U}^{-1} \right)_{kl} &= -E\left[\frac{\partial^2 \ln L}{\partial \mu_k \partial \mu_l} \right] \\ &= \sum_{i=1}^{N} \frac{E[n_i] R_{ik} R_{il}}{\nu_i^2} \\ &= \sum_{i=1}^{N} \frac{R_{ik} R_{il}}{\nu_i}. \end{aligned} \tag{16.2.13}$$

等式两边同时乘一次 $\underset{\sim}{U}$, 两次 $\underset{\sim}{R}^{-1}$, 并对适当的下标求和, 可由式 (16.2.13) 解出协方差矩阵

$$U_{ij} = \sum_{k=1}^{N} \left(\underset{\sim}{R}^{-1} \right)_{ik} \left(\underset{\sim}{R}^{-1} \right)_{jk} \nu_k. \tag{16.2.14}$$

这与严格计算导出的式 (16.2.7) 结果相同, 所以我们看到极大似然解是无偏、有效估计量, 也即对于无偏估计量, 它具有最小的可能方差. 利用最小二乘法, 我们将得到同样的结果; 这种情形下, 高斯–马尔可夫原理保证了最小二乘估计量为无偏、有效估计量.

尽管图 16.1(d) 所表示的解与真值分布很少有类似之处, 但它仍然有一些可取的性质. 它的构建简单, 偏差为 0, 其方差等于最小方差界. 但是为了使得它能够付诸实用, 必须考虑不同子区间估计量 $\hat{\mu}_j$ 的关联. 例如我们可以通过构建一个 χ^2 统计量

$$\chi^2 = (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)^T \underset{\sim}{U}^{-1} (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0) \tag{16.2.15}$$

来检验估计量 $\hat{\boldsymbol{\mu}}$ 与假设 $\boldsymbol{\mu}_0$ 的一致性, 其中使用了估计量的完整的协方差矩阵 $\underset{\sim}{U}$. 如果 χ^2 只用 $\underset{\sim}{U}$ 的对角元素计算, 这一检验将毫无意义. 我们还应当指出, 尽管这

一例子中方差很大, 如果将子区间取得比分辨函数的宽度要大, 则方差能够明显地减小.

抛开其缺点不说, 响应矩阵求逆方法给我们指出了某些重要的经验, 并为实际可使用的更好的去弥散方法提供了一个出发点. 由于逆矩阵解具有 0 偏差和最小方差, 故可知, 要想进一步减小估计量的方差只可能通过引入某种偏差才能达到. 去弥散的技巧在于以某种方式构建一个有偏估计量 $\hat{\boldsymbol{\mu}}$, 当构建该估计量时需要引入关于平滑性假设的先验知识, 如果这一先验假设是正确的, 则该偏差将是个小量. 粗略地讲, 目的是寻找估计量的偏差和方差之间的一种最优平衡. 我们将在 16.7 节看到, 在确定怎样达到这一最优平衡方面存在一定的任意性.

加入先验知识的需要使得我们要利用贝叶斯方法, 它将先验概率与数据结合起来产生关于真值分布的后验概率 (见 13.2 节), 这是关于去弥散方法的文献中的一个公共出发点. 然而, 它会遇到下述困难：先验知识往往是复杂和定性的, 难以表示为先验概率. 先验信度本质上带有主观性这一事实在这里并不是一种真正的缺点; 在经典方法中, 怎样选择一种有偏估计量同样带有一定的主观性. 下面我们将主要按照经典统计方法, 利用估计量的偏差和方差作为判断去弥散方法对于问题的解的品质的判据, 同时指出与贝叶斯方法的相互联系.

最后, 我们来考虑真实直方图子区间数 M 和观测直方图子区间数 N 不相同的情形. 对 $M > N$, 式 (16.1.12) $\boldsymbol{\nu} = \underset{\sim}{R}\,\boldsymbol{\mu} + \boldsymbol{\beta}$ 是欠定 (underdetermined) 方程组, 其解不唯一. 16.4 节所陈述的正规化去弥散方法可给出估计量 $\hat{\boldsymbol{\mu}}$ 的解. 对 $M < N$, 方程组 (16.1.12) 是超定的, 一般不存在严格解. 其近似解可利用极大似然法或最小二乘法来构建, 即问题与第 7 到第 10 章讨论的参数估计等价. 如果 M 很大, 则估计量间的关联可导致大的方差. 这种情形下, 可能需要利用 16.4 节所述的正规化方法, 以引入偏差作为代价来减小方差.

16.3 修正因子法

考虑真值分布直方图 $\boldsymbol{\mu}$ 子区间数和观测数据直方图 $\boldsymbol{n}$ 子区间数相同的情形. 一种最简单、最常用的方法是将 μ_i 的估计量取为

$$\hat{\mu}_i = C_i\,(n_i - \beta_i)\,, \tag{16.3.1}$$

式中 β_i 是本底的期望值, C_i 是一**修正因子**. 利用蒙特卡罗方法模拟所研究的过程以及测量装置, 可以确定该修正因子. 分别运行包含和不包含测量装置模拟的蒙特卡罗程序, 可分别得出每个子区间中实验观测事例数和真实事例数的模型预期值 ν_i^{MC} 和 μ_i^{MC}, 从而确定修正因子 C_i. 这里 ν^{MC} 仅指信号过程, 本底不包含在内.

修正因子直接由两者之比值求得

$$C_i = \frac{\mu_i^{MC}}{\nu_i^{MC}}. \tag{16.3.2}$$

我们假定, 能够产生足够数量的蒙特卡罗数据, 使修正因子的统计误差可以忽略. 如果情况不是如此, 则 C_i 的不确定性可以通过通常的误差传递合并到估计值 $\hat{\mu}_i$ 的误差中.

如果分辨的效应可以忽略, 则响应矩阵是对角化的, 即 $R_{ij} = \delta_{ij}\varepsilon_j$, 因而可有

$$\nu_i^{\text{sig}} = \nu_i - \beta_i = \varepsilon_i \mu_i, \tag{16.3.3}$$

式中 ν_i^{sig} 是无本底情形下子区间 i 中的预期观测事例数. 于是修正因子变成 $C_i = 1/\varepsilon_i$, 这样 $1/C_i$ 起到一个总体效率的作用. 但是, 当响应矩阵包含非对角项, $1/C_i$ 的值可能大于 1. 也就是说由于不同子区间的事例之间的相互迁移, 在一给定的子区间中出现的观测事例数可能多于真值分布在此产生的事例数.

估计量 $\hat{\mu}_i$ 的期望值是

$$\begin{aligned} E\left(\hat{\mu}_i\right) &= C_i E\left(n_i - \beta_i\right) = C_i\left(\nu_i - \beta_i\right) = \frac{\mu_i^{MC}}{\nu_i^{MC}} \nu_i^{\text{sig}} \\ &= \left(\frac{\mu_i^{MC}}{\nu_i^{MC}} - \frac{\mu_i}{\nu_i^{\text{sig}}}\right) \nu_i^{\text{sig}} + \mu_i. \end{aligned} \tag{16.3.4}$$

因此估计量 $\hat{\mu}_i$ 是有偏的, 其偏差由上式右边的第一项表示; 仅当比值 μ_i/ν_i^{sig} 对于蒙特卡罗模型和实验数据是相等的情形下, 该偏差才为 0.

估计量的协方差矩阵由下式给定

$$\begin{aligned} \operatorname{cov}\left(\hat{\mu}_i, \hat{\mu}_j\right) &= C_i^2 \operatorname{cov}\left(n_i, n_j\right) \\ &= C_i^2 \delta_{ij} \nu_i. \end{aligned} \tag{16.3.5}$$

上式的最后结果利用了 n_i 是期望值为 ν_i、相互独立的泊松变量情形下的协方差矩阵的性质. 在许多实际问题中, C_i 值的量级为 1, 因此估计量 $\hat{\mu}_i$ 的方差与理想分辨情形下估计量 $\hat{\mu}_i$ 的方差相近. 此外, 该方法容易实现, 甚至不需要矩阵求逆. 所付出的代价是引入了偏差

$$b_i = \left(\frac{\mu_i^{MC}}{\nu_i^{MC}} - \frac{\mu_i}{\nu_i^{\text{sig}}}\right) \nu_i^{\text{sig}}. \tag{16.3.6}$$

这一偏差导致的系统不确定性可以用模拟所研究过程和探测器的不同模型计算得到的修正因子的差别来估计. 显然, 一个较好的模型对应于一个较小的偏差, 因此通常建议蒙特卡罗模型可以通过所得到的估计分布 $\hat{\boldsymbol{\mu}}$ 来进行调整, 亦即通过调节

蒙特卡罗模型参数来改善 $\boldsymbol{\mu}^{MC}$ 与扣除了本底的数据 $\boldsymbol{n}-\boldsymbol{\beta}$ 之间的一致性. 迭代地实施这种调整以从调整后的模型求得改善了的修正因子.

修正因子法的危险性在于, 偏差通常将估计量 $\hat{\boldsymbol{\mu}}$ 拉向模型的预期值 $\boldsymbol{\mu}^{MC}$. 这会造成模型检验这一任务的复杂化, 而检验模型可能是进行实验测量的目的之一. 这种情形下, 我们必须保证, 去弥散结果中由于修正因子的模型依赖导致的不确定性在估计系统误差时已考虑在内, 并保证这些因素在模型检验中已考虑在内.

16.4　正规化去弥散的一般策略

虽然修正因子法简单且在实际中广泛采用, 但它有若干缺点, 主要是其结果具有模型依赖性. 去弥散另一种可选的途径是要求真值直方图 $\boldsymbol{\mu}$ 的估计量具有平滑性. 这种方法称为去弥散分布的**正规化** (regularization).

我们首先回到 16.2 节所述的由响应矩阵求逆得到的振荡解的情形. 对于真值直方图 $\boldsymbol{\mu}$ 的这一估计是使对数似然函数达到极大值 $\ln L_{\max}$, 或者是 χ^2 达到极小求得的. 下面, 我们只讨论极大似然函数的情形; 利用 χ^2 方法的对应关系可通过 $\ln L=-\chi^2/2$ 的替代来求得.

我们将 $\boldsymbol{\mu}-$ 空间中极大似然 (或最小二乘) 解附近的某个确定的区域考虑为可接受的解区, 即在此区域内测量值的期望值 $\boldsymbol{\nu}$ 与数据 $\boldsymbol{n}$ 之间的一致性达到了可接受的水平. 该区域的大小由区域内的 $\ln L$ 与 $\ln L_{\max}$ 之差小于某个确定值来决定. 也就是说, $\boldsymbol{\mu}-$ 空间中的可接受的解区由 (极大似然法)

$$\ln L(\boldsymbol{\mu}) \geqslant \ln L_{\max}-\Delta \ln L \tag{16.4.1}$$

或对于最小二乘法由

$$\chi^2(\boldsymbol{\mu}) \leqslant \chi^2_{\min}+\Delta\chi^2 \tag{16.4.2}$$

来确定, 其中 $\Delta\ln L$ 或 $\Delta\chi^2$ 的数值需要适当地选择. $\Delta\ln L$ 或 $\Delta\chi^2$ 的数值将决定去弥散后的直方图的偏差和方差之间的平衡, 16.6 节将对此进行详细讨论.

除了可接受解区之外, 我们还需要引入一个称为**正规函数**的函数 $S(\boldsymbol{\mu})$ 来表征解的平滑性程度. $S(\boldsymbol{\mu})$ 的若干可能的形式将在下一节讨论. 一般的策略是在由式 (16.4.1) 或式 (16.4.2) 所确定的可接受解中选择平滑度最高的那个解.

在 $\ln L(\boldsymbol{\mu})=\ln L_{\max}-\Delta\ln L$ 的约束下求正规函数 $S(\boldsymbol{\mu})$ 的极大, 等价于量

$$\alpha\left[\ln L(\boldsymbol{\mu})-(\ln L_{\max}-\Delta\ln L)\right]+S(\boldsymbol{\mu}) \tag{16.4.3}$$

同时对 $\boldsymbol{\mu}$ 和 α 求极大. 这里 α 是拉格朗日乘子, 称为**正规化参数**, 它的数值与一特定的 $\Delta\ln L$ 值相对应. 对一给定的 α, 可通过找出 $\ln L$ 和 $S(\mu)$ 的加权和

$$\Phi(\boldsymbol{\mu})=\alpha\ln L(\boldsymbol{\mu})+S(\boldsymbol{\mu}) \tag{16.4.4}$$

的极大值来求得问题的解. α 设置为 0 对应于最平滑的分布; 这相当于完全忽略数据的作用. 当 α 很大时与似然函数达到极大相对应, 其解对应于由响应矩阵求逆得到的振荡解.

为了使 $\Phi(\boldsymbol{\mu})$ 求极大得到的解在事实上等价于前面所述的一般策略, 即等于式 (16.4.1) 或式 (16.4.2) 所确定的可接受解中平滑度最高的那个解, $\ln L(\boldsymbol{\mu})$ 等于常数的表面和函数 $S(\boldsymbol{\mu})$ 的行为必须足够良性, 下面的讨论中我们就做了这样的假定. 特别是, 这些表面不应既有凸面又有凹面, 或者形态非常复杂, 以至于存在多个局域极大.

因为关系式 $\boldsymbol{\nu} = \underset{\sim}{R}\boldsymbol{\mu} + \boldsymbol{\beta}$ 成立, 我们能够将 L 和 S 视为 $\boldsymbol{\mu}$ 或 $\boldsymbol{\nu}$ 的函数. 与此类似, 我们将总是利用关系式

$$\hat{\boldsymbol{\nu}} = \underset{\sim}{R}\hat{\boldsymbol{\mu}} + \boldsymbol{\beta} \tag{16.4.5}$$

来确定 $\hat{\boldsymbol{\nu}}$ 的估计量; 而知道了这些关系等价于知道了估计量 $\hat{\boldsymbol{\mu}}$. 但需要指出, 与 16.2 节所述的响应矩阵求逆方法不同, 我们将不再有关系式 (16.2.2) $\hat{\boldsymbol{\nu}} = \boldsymbol{n}$. 还应该记住, $\mu_{\text{tot}} = \sum_j \mu_j$ 和 $\nu_{\text{tot}} = \sum_i \nu_i = \sum_{i,j} R_{ij}\mu_j$ 也是 $\boldsymbol{\mu}$ 的函数.

这里我们将只考虑事例的估计总数 $\hat{\nu}_{\text{tot}}$ 等于实际观测数 n_{tot} 情形下的估计量 $\hat{\boldsymbol{\mu}}$:

$$\hat{\nu}_{\text{tot}} = \sum_{i=1}^{N} \hat{\nu}_i = \sum_{i=1}^{N}\sum_{j=1}^{M} R_{ij}\hat{\mu}_j + \beta_i = n_{\text{tot}}. \tag{16.4.6}$$

这一条件一般情形下并不自动满足. 将式 (16.4.4) 修改为

$$\varphi(\boldsymbol{\mu}, \lambda) = \alpha \ln L(\boldsymbol{\mu}) + S(\boldsymbol{\mu}) + \lambda\left[n_{\text{tot}} - \sum_{i=1}^{N}\nu_i\right], \tag{16.4.7}$$

其中 λ 是拉格朗日乘子, 并令 $\partial\varphi/\partial\lambda = 0$, 就得到 $\sum_i \nu_i = n_{\text{tot}}$, 从而条件式 (16.4.6) 得以满足.

应当指出, 正规化参数 α 放在正规函数 $S(\boldsymbol{\mu})$ 之前 (许多文献中采用这种写法) 还是似然函数之前对结果没有区别.

16.5 正 规 函 数

16.5.1 Tikhonov 正规函数

一种广泛应用的衡量平滑性的尺度是真值分布的 k 阶导数的平方平均值. 这种方法是由 Phillips[149] 和 Tikhonov[156,157] 各自独立地提出的, 通常称为**Tikhonov 正规化**. 如果我们考虑离散化为直方图之前的概率密度函数 $f_{\text{true}}(y)$, 则正规函数

是

$$S\left[f_{\text{true}}\left(y\right)\right]=-\int\left(\frac{d^{k}f_{\text{true}}\left(y\right)}{dy^{k}}\right)^{2}\mathrm{d}y,\tag{16.5.1}$$

式中积分遍及 y 的一切容许值. 式中的负号是由于我们约定对式 (16.4.7) 定义的 φ 需要求极大. S 越大, 平滑度越高. 等价地, 也可以对正规函数和对数似然函数的组合加上负号后求极小, 这种约定在文献中也常常遇到.

原则上, 可以使用不同阶导数项的线性组合; 而在实际情形中, 通常选择一定阶的导数项, 即一个 k 值. 当 $f_{\text{true}}\left(y\right)$ 用一直方图来表示时, 导数用有限差分代替. 对于各子区间等宽的情形, 当 $k=1$, 可以利用 [150]

$$S\left(\boldsymbol{\mu}\right)=-\sum_{i=1}^{M-1}\left(\mu_{i}-\mu_{i+1}\right)^{2},\tag{16.5.2}$$

当 $k=2$, 则

$$S\left(\boldsymbol{\mu}\right)=-\sum_{i=1}^{M-2}\left(-\mu_{i}+2\mu_{i+1}-\mu_{i+2}\right)^{2},\tag{16.5.3}$$

当 $k=3$, 则

$$S\left(\boldsymbol{\mu}\right)=-\sum_{i=1}^{M-3}\left(-\mu_{i}+3\mu_{i+1}-3\mu_{i+2}+\mu_{i+3}\right)^{2}.\tag{16.5.4}$$

对于导数阶数的常用选择是 $k=2$, 这时 $S\left(\boldsymbol{\mu}\right)$ 与平均曲率相关联.

如果各子区间宽度 Δy_i 均相等, 则它们可以略而不计, 如在式 (16.5.2)$\sim$ 式 (16.5.4) 中的情形. 子区间宽度的效应可以吸收到正规化参数 α 之中. 如果各子区间宽度 Δy_i 并不全部相等, 那么它们可以直接吸收到有限差分之中. 例如对 $k=2$, 我们可以认为在每一组 3 个相邻的子区间中 $f_{\text{true}}\left(y\right)$ 有抛物线的形式：

$$f_{i}\left(y\right)=a_{0i}+a_{1i}y+a_{2i}y^{2}.\tag{16.5.5}$$

共有 $M-2$ 个这样的组, 处于中心的子区间编号为 $i=2,\cdots,M-1$. 每一组中的 3 个系数 a_{0i},a_{1i},a_{2i} 可由 $f_i\left(y\right)$ 在子区间 $i-1,i,i+1$ 中的积分值分别等于 $\mu_{i-1},\mu_i,\mu_{i+1}$ 来确定. 中心位于子区间 i 的那一组其二阶导数为 $f_i''=2a_{2i}$, 于是正规函数的形式为

$$S\left(\boldsymbol{\mu}\right)=-\sum_{i=2}^{M-1}f_{i}''^{2}\Delta y_{i}.\tag{16.5.6}$$

需要指出, 在第一和最后一个子区间, 二阶导数无法确定. 这里, 它们没有包含在求和式 (16.5.6) 之中, 即它们被取为 0; 另一种做法是它们可设置为等于子区间 2 和 $M-1$ 的二阶导数值.

不论取何种 k 值, 也不论子区间取何种宽度, 上述 $S(\boldsymbol{\mu})$ 均可表示为如下形式:

$$S(\boldsymbol{\mu}) = -\sum_{i,j=1}^{M} G_{ij}\mu_i\mu_j = -\boldsymbol{\mu}^T \underset{\sim}{G}\,\boldsymbol{\mu}, \tag{16.5.7}$$

其中 $\underset{\sim}{G}$ 是一个对称常数矩阵. 例如, 对于 $k = 2$ 且各子区间等宽的正规函数式 (16.5.3), $\underset{\sim}{G}$ 由下式给定:

$$\left.\begin{aligned} &G_{ii} = 6 \\ &G_{i,i\pm1} = G_{i\pm1,i} = -4 \\ &G_{i,i\pm2} = G_{i\pm2,i} = 1 \end{aligned}\right\} \quad 3 \leqslant i \leqslant M-2,$$

$$G_{11} = G_{MM} = 1, \tag{16.5.8}$$

$$G_{22} = G_{M-1,M-1} = 5,$$

$$G_{12} = G_{21} = G_{M,M-1} = G_{M-1,M} = -2,$$

G_{ij} 的其他各项均为 0.

为了求得估计量及其协方差矩阵 (见 16.6 节), 我们需要知道 S 的一阶和二阶导数. 它们是

$$\frac{\partial S}{\partial \mu_i} = -2\sum_{j=1}^{M} G_{ij}\mu_j \tag{16.5.9}$$

和

$$\frac{\partial^2 S}{\partial \mu_i \partial \mu_j} = -2G_{ij}. \tag{16.5.10}$$

$k = 2$ 的 Tikhonov 正规函数广泛地应用于粒子物理中结构函数 (轻子–核子散射中的运动学变量分布) 的去弥散计算. 详细的介绍可参见文献 [139], [143], [151], [159].

16.5.2 基于极大熵原理的正规函数

另一种广泛使用的正规函数是基于某个概率分布 $\boldsymbol{p} = (p_1, \cdots, p_M)$ 的**熵** H, 它的定义是 [152]

$$H = -\sum_{i=1}^{M} p_i \ln p_i. \tag{16.5.11}$$

这里的基本思想是将熵作为直方图 $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_M)$ 平滑性的尺度, 并利用

$$S(\boldsymbol{\mu}) = H(\boldsymbol{\mu}) = -\sum_{i=1}^{M} \frac{\mu_i}{\mu_{\text{tot}}} \ln \frac{\mu_i}{\mu_{\text{tot}}} \tag{16.5.12}$$

作为正规函数. 基于式 (16.5.12) 的正规函数 $S(\boldsymbol{\mu})$ 构建的估计量 $\hat{\mu}$ 称为基于**极大熵原理**的估计量. 为了说明熵与平滑度之间的关系, 我们来考察总事例数为 μ_{tot} 时, 构造出某一特定直方图 $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_M)$ 可以有多少种方式 (这里 μ_j 是整数). 答案是

$$\varOmega(\mu) = \frac{\mu_{\text{tot}}!}{\mu_1!\mu_2!\cdots\mu_M!}. \tag{16.5.13}$$

(多项分布的概率表示式 (4.2.1) 中出现类似的因子). 对式 (16.5.13) 取对数并应用斯特林近似 $\ln n! \approx n(\ln n - 1)$(大 n 值该近似成立), 立即可得

$$\begin{aligned}\ln \varOmega &\approx \mu_{\text{tot}}(\ln \mu_{\text{tot}} - 1) - \sum_{i=1}^{M} \mu_i (\ln \mu_i - 1) \\ &= -\sum_{i=1}^{M} \mu_i \ln \frac{\mu_i}{\mu_{\text{tot}}} \\ &= \mu_{\text{tot}} H(\mu) = \mu_{\text{tot}} S(\mu).\end{aligned} \tag{16.5.14}$$

我们将利用方程式 (16.5.14) 把 $\ln \varOmega$ 推广到 μ_i 不是整数的情形.

如果所有事例集中在单个子区间, 则只有一种安排方式 $\varOmega = 1$, 这时直方图具有最小的平滑度, 于是熵达到极小 $S(\boldsymbol{\mu}) = H(\boldsymbol{\mu}) = 0$. 另一种极端情况是, 所有的 μ_i 相等, 即直方图对应于均匀分布, 则可以证明熵达到极大. 为了在约束条件 $\sum_i p_i = 1$ 下求 H 的极大, 可以利用拉格朗日乘子法.

为了后面的使用方便, 这里列出熵正规函数 $S(\boldsymbol{\mu})$ 的一阶和二阶导数:

$$\frac{\partial S}{\partial \mu_i} = -\frac{1}{\mu_{\text{tot}}} \ln \frac{\mu_i}{\mu_{\text{tot}}} - \frac{S(\mu)}{\boldsymbol{\mu}_{\text{tot}}} \tag{16.5.15}$$

以及

$$\frac{\partial^2 S}{\partial \mu_i \partial \mu_j} = -\frac{1}{\mu_{\text{tot}}^2}\left[1 - \frac{\delta_{ij}\mu_{\text{tot}}}{\mu_i} + \ln\left(\frac{\mu_i \mu_j}{\mu_{\text{tot}}^2}\right) + 2S(\boldsymbol{\mu})\right]. \tag{16.5.16}$$

16.5.3 贝叶斯统计的极大熵原理

在许多关于去弥散问题的文献中, 极大熵原理是在贝叶斯统计架构中发展起来的 [85,145,150]. 然而, 贝叶斯统计的极大熵原理应用于去弥散问题会陷入困境, 这一点我们会在本节后面以及 16.8 节加以讨论. 不过, 将贝叶斯的极大熵原理与前一节的经典方法的极大熵原理加以对比, 还是很有意义的.

贝叶斯方法中, $\boldsymbol{\mu}$ 处理为主观概率意义上的随机变量 (见 13.2 节), 联合概率密度 $f(\boldsymbol{\mu}|\boldsymbol{n})$ 表示真值直方图是由 $\boldsymbol{\mu}$ 生成的信度. 为了依据观测数据 $\boldsymbol{n}$ 来更新我们关于 $\boldsymbol{\mu}$ 的知识, 我们利用贝叶斯定理

$$f(\boldsymbol{\mu}|\boldsymbol{n}) \propto L(\boldsymbol{n}|\boldsymbol{\mu})\,\pi(\boldsymbol{\mu}), \tag{16.5.17}$$

其中 $L(\boldsymbol{n}|\boldsymbol{\mu})$ 是似然函数 (给定 $\boldsymbol{\mu}$ 的条件下观测数据为 $\boldsymbol{n}$ 的条件概率), $\pi(\boldsymbol{\mu})$ 是先验密度. 先验密度表示我们在收集观测数据 $\boldsymbol{n}$ 之前的关于真值 $\boldsymbol{\mu}$ 的知识.

这里, 我们将事例总数 μ_{tot} 视为整数. 这与经典方法中不同, 在那里 μ_{tot} 表示的是一个整数随机变量的期望值, 因此它本身不一定是一个整数. 假定我们对这 μ_{tot} 个事例在直方图中的分布没有先验知识. 按照对称性要求, 可以认为这 μ_{tot} 个事例分配在直方图的 M 个子区间中的每一种可能方式具有相等的可能性. 因此, 在没有任何其他先验信息的情形下, 某一特定直方图 $(\mu_1, \cdots, \mu_M)$ 出现的概率应当正比于总事例数为 μ_{tot} 时构造出该直方图的方式总数; 后者即是方程式 (16.5.13) 给定的 Ω. 方式总数 $\Omega(\boldsymbol{\mu})$ 于是被解释为先验概率 $\pi(\boldsymbol{\mu})$, 由式 (16.5.14) 立即有

$$\begin{aligned}\pi(\boldsymbol{\mu}) = \Omega(\boldsymbol{\mu}) &= \frac{\mu_{\text{tot}}!}{\mu_1!\mu_2!\cdots\mu_M!} \\ &= \exp(\mu_{\text{tot}} H),\end{aligned} \tag{16.5.18}$$

式中, 熵 H 由方程式 (16.5.11) 定义.

从贝叶斯统计观点而言, 当我们确定了联合概率密度 $f(\boldsymbol{\mu}|\boldsymbol{n})$, 问题就解决了. 然而, 完整地描述 $f(\boldsymbol{\mu}|\boldsymbol{n})$ 是不实际的, 因为这是 M 个变量的函数, M 等于去弥散分布的子区间个数. 因此必须找到某种综合这些变量信息的方法, 典型的做法是选择单个向量 $\hat{\boldsymbol{\mu}}$ 作为贝叶斯估计量, 通常是选择使得概率 $f(\boldsymbol{\mu}|\boldsymbol{n})$ 或它的对数达到极大的那个 $\boldsymbol{\mu}$ 值. 按照方程式 (16.5.17), 这可以通过对函数

$$\begin{aligned}\ln f(\boldsymbol{\mu}|\boldsymbol{n}) &\propto \ln L(\boldsymbol{\mu}|\boldsymbol{n}) + \ln \pi(\boldsymbol{\mu}) \\ &= \ln L(\boldsymbol{\mu}|\boldsymbol{n}) + \boldsymbol{\mu}_{\text{tot}} H(\boldsymbol{\mu})\end{aligned} \tag{16.5.19}$$

求极大求得. 与正规化去弥散方法的一般表示式 (16.4.4) 对比, 可知贝叶斯极大熵原理对应于利用正规函数

$$S(\boldsymbol{\mu}) = \mu_{\text{tot}} H(\boldsymbol{\mu}) = -\sum_{i=1}^{M} \mu_i \ln \frac{\mu_i}{\mu_{\text{tot}}}. \tag{16.5.20}$$

同时, 正规化参数 α 这时不再是一个任意因子, 而需设定为 1. 如果所有的效率 ε_i 相等, 则要求 $\nu_{\text{tot}} = n_{\text{tot}}$ 就意味着 μ_{tot} 是一个常数. 这等价于利用经典极大熵原理的正规函数 $S(\boldsymbol{\mu}) = H$ 而 $\alpha = 1/\mu_{\text{tot}}$.

但是当效率 ε_i 并不全都相等, 则 ν_{tot} 为常数并不表示 μ_{tot} 是一个常数, 其结果是 $S(\boldsymbol{\mu}) = \mu_{\text{tot}} H(\boldsymbol{\mu})$ 熵极大的 $\boldsymbol{\mu}$ 分布不再是均匀分布. 这是因为 S 会随着 μ_{tot} 的增加而增加, 于是在熵 S 极大的 $\boldsymbol{\mu}$ 分布中, 效率低的子区间的值会升高. 这种情形下, 利用 H 和 $\mu_{\text{tot}} H$ 作为正规函数会导致略微不同的结果, 虽然当效率 ε_i 具有相同量级的情形下结果的差别并不大. 无论如何, $S = H$ 作为表示平滑性的一种尺度是恰当的, 因为 H 极大的 $\boldsymbol{\mu}$ 分布总是均匀的.

在 16.8 节中我们将看到, 对式 (16.5.19) 的函数求极大得到的贝叶斯估计量对于包含熵的项赋予了过多的权值 (见图 16.3(a) 以及参考文献 [155]). 从经典统计的观点来看, 我们可以认为它没有达到估计量的偏差和方差之间的一种好的平衡, 而是具有不合理的过大的偏差. 作为贝叶斯方案的修正形式, 我们可以对式 (16.5.19) 中的 μ_{tot} 用一个有效事例数 μ_{eff} 代替, 它可以调节到小于 μ_{tot}. 于是估计量由

$$\ln L(\boldsymbol{\mu}|\boldsymbol{n}) + \mu_{\text{eff}} H(\boldsymbol{\mu}) \tag{16.5.21}$$

的极大值决定. 这等价于经典方法极大熵原理的正规函数 $S(\boldsymbol{\mu}) = H(\boldsymbol{\mu})$, 而参数 μ_{eff} 起到了正规化参数的作用.

贝叶斯解存在的问题来源于我们将 $\Omega(\boldsymbol{\mu})$ 用作先验密度. 不论是从贝叶斯统计还是经典统计的观点出发, 量 $\boldsymbol{p} = \boldsymbol{\mu}/\mu_{\text{tot}}$ 总是由一组未知常数所给定, 例如, β 衰变的例子中 $\boldsymbol{p}$ 由电子能量分布给定. 不论是哪一种统计观点, 我们关于 $\boldsymbol{p}$ (不是 $\boldsymbol{\mu}$) 的分布**形式**的先验知识应当独立于我们获得的数据样本的观测个数. 但这一点使得将 $\Omega(\boldsymbol{\mu})$ 视为先验密度 $\pi(\boldsymbol{\mu})$ 会产生根本性的问题, 因为当 μ_{tot} 增加时, $\pi(\boldsymbol{\mu}) = \Omega(\boldsymbol{\mu})$ 会进一步倾向于均匀分布 (即所有的 p_i 相等).

通常的情况是, 我们可能具有关于 $\boldsymbol{p}$ 的分布形式的某种先验知识, 但是难以定量化. 例如, 我们可以预期这种分布先验地很不可能有大量的结构, 因为很难想象一种物理理论会预期该分布具有许多个峰值. 另一方面, 一个完全平坦的分布看起来也不合乎物理要求, 所以 $\Omega(\boldsymbol{\mu})$ 也确实不能反映我们的先验知识. 由于将 $\Omega(\boldsymbol{\mu})$ 解释为先验概率密度存在这些困难, 这里我们将采用经典的观点, 只是将熵视为一种可能的正规函数, 即认为 $S(\boldsymbol{\mu}) = H(\boldsymbol{\mu})$.

16.5.4　基于交叉熵的正规函数

在 16.5.2 节中已经指出, 当熵达到极大时直方图对应于平分布, 因此将极大熵原理引入 $\boldsymbol{\mu}$ 的估计量中将使去弥散的结果拉向较为均匀的分布. 假定我们已知一种分布 $\boldsymbol{q} = (q_1, \cdots, q_M)$, 认为它非常像是真值分布 $\boldsymbol{p} = \boldsymbol{\mu}/\boldsymbol{\mu}_{\text{tot}}$ 的先验形状. 我们称 $\boldsymbol{q}$ 为**参考分布**. 但是, 假设我们确实不知道怎样将我们对于 q 的先验信度定量化, 因此我们没有应用贝叶斯原理所需的先验密度 $\pi(\boldsymbol{\mu})$ 可利用. 也就是说, $\boldsymbol{q}$ 表示的是先验密度达到极大情形下的归一化直方图 $\boldsymbol{\mu}/\boldsymbol{\mu}_{\text{tot}}$, 但是它并不能完整地描述先验密度.

这种情形下, 正规函数可取为

$$S(\boldsymbol{\mu}) = K(\boldsymbol{p}; \boldsymbol{q}), \tag{16.5.22}$$

式中 $K(\boldsymbol{p}; \boldsymbol{q})$ 称为**交叉熵**[146] (cross-entropy) 或**Shannon-Jaynes 熵**[144], 其定义

为

$$K(\boldsymbol{p};\boldsymbol{q}) = -\sum_{i=1}^{M} p_i \ln \frac{p_i}{Mq_i}. \tag{16.5.23}$$

通常的交叉熵定义中对数项不含 M, 求和号前也没有负号, 在这样的通常的交叉熵定义下极大熵原理变成为极小交叉熵原理. 我们的式 (16.5.23) 定义中将保留负号, 以保证 $K(\boldsymbol{p};\boldsymbol{q})$ 与式 (16.5.11) 所示的 Shannon 熵 $H(\boldsymbol{p})$ 之间的相似性. 应当指出, 当参考分布为均匀分布, 即对一切 i 有 $q_i = 1/M$ 时, $K(\boldsymbol{p};\boldsymbol{q}) = H(\boldsymbol{p})$.

容易证明, 当概率 $\boldsymbol{p}$ 等于参考分布 $\boldsymbol{q}$ 的概率时, 交叉熵 $K(\boldsymbol{p};\boldsymbol{q})$ 达到极大. 利用式 (16.5.22) 的正规函数的效果是, 如果真值分布等于 (或接近于) 参考分布, 则估计量 $\hat{\boldsymbol{\mu}}$ 的偏差将等于 0(或接近于 0).

16.6　估计量的方差和偏差

估计量 $\hat{\boldsymbol{\mu}}$ 是观测数据 $\boldsymbol{n}$ 的函数, 因而它自身亦为随机变量. 为了求得协方差矩阵 $U_{ij} = \mathrm{cov}(\hat{\mu}_i, \hat{\mu}_j)$, 我们需要求得 $\hat{\boldsymbol{\mu}}$ 作为 $\boldsymbol{n}$ 的函数的一个近似表式, 然后利用误差传递公式求得估计量 $\hat{\boldsymbol{\mu}}$ 的协方差矩阵 $\underset{\sim}{U}$ 与数据 $\boldsymbol{n}$ 的协方差矩阵 $V_{ij} = \mathrm{cov}(n_i, n_j)$ 之间的相互关系.

估计量 $\hat{\boldsymbol{\mu}}$ 通过对式 (16.4.7) 给定的函数 $\varphi(\boldsymbol{\mu}, \lambda)$ 求极大得到, $\varphi(\boldsymbol{\mu}, \lambda)$ 包含对数似然函数或 χ^2 函数, 以及某种形式的正规函数 $S(\boldsymbol{\mu})$ (Tikhonov 正规函数, 熵, 交叉熵等). 估计量 $\hat{\boldsymbol{\mu}}$ 和拉格朗日乘子 λ 是 M+1 个方程的解:

$$F_i(\boldsymbol{\mu}, \lambda, \boldsymbol{n}) = 0, \quad i = 1, \cdots, M+1, \tag{16.6.1}$$

其中,

$$F_i(\boldsymbol{\mu}, \lambda, \boldsymbol{n}) = \begin{cases} \dfrac{\partial \varphi}{\partial \mu_i}, & i = 1, \cdots, M, \\ \dfrac{\partial \varphi}{\partial \lambda}, & i = M+1. \end{cases} \tag{16.6.2}$$

设实际得到的数据由向量 $\tilde{\boldsymbol{n}}$ 给定, 相应的估计值为 $\tilde{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}(\tilde{\boldsymbol{n}})$, 拉格朗日乘子 λ 的值为 $\tilde{\lambda}$. 我们希望知道, 如果观测数据由某个不同的 $\boldsymbol{n}$ 值给定, $\hat{\boldsymbol{\mu}}$ 和 λ 将如何变化. 将函数 $F_i(\boldsymbol{\mu}, \lambda, \boldsymbol{n})$ 在 $\tilde{\boldsymbol{\mu}}$, $\tilde{\lambda}$ 和 $\tilde{\boldsymbol{n}}$ 值邻域做一阶泰勒展开可得

$$\begin{aligned} F_i(\boldsymbol{\mu}, \lambda, \boldsymbol{n}) \approx & F_i\left(\tilde{\boldsymbol{\mu}}, \tilde{\lambda}, \tilde{\boldsymbol{n}}\right) + \sum_{j=1}^{M} \left[\frac{\partial F_i}{\partial \mu_j}\right]_{\tilde{\boldsymbol{\mu}}, \tilde{\lambda}, \tilde{\boldsymbol{n}}} (\mu_j - \tilde{\mu}_j) \\ & + \left[\frac{\partial F_i}{\partial \lambda}\right]_{\tilde{\boldsymbol{\mu}}, \tilde{\lambda}, \tilde{\boldsymbol{n}}} \left(\lambda - \tilde{\lambda}\right) + \sum_{j=1}^{N} \left[\frac{\partial F_i}{\partial n_j}\right]_{\tilde{\boldsymbol{\mu}}, \tilde{\lambda}, \tilde{\boldsymbol{n}}} (n_j - \tilde{n}_j). \end{aligned} \tag{16.6.3}$$

上式的左边 $F_i(\boldsymbol{\mu},\lambda,\boldsymbol{n})$ 及右边第一项 $F_i\left(\tilde{\boldsymbol{\mu}},\tilde{\lambda},\tilde{\boldsymbol{n}}\right)$ 均等于 0, 因为它们各自的一组自变量都是方程的一组解. 方程式 (16.6.3) 对 $\boldsymbol{\mu}$ 求解得到

$$\hat{\boldsymbol{\mu}}(\boldsymbol{n}) \approx \tilde{\boldsymbol{\mu}} - \underset{\sim}{A}^{-1}\underset{\sim}{B}(\boldsymbol{n}-\tilde{\boldsymbol{n}}), \tag{16.6.4}$$

式中 $\boldsymbol{\mu}$ 的第 $M{+}1$ 分量代表拉格朗日乘子 λ. $(M+1)\times(M+1)$ 阶对称矩阵 $\underset{\sim}{A}$ 由下式给定：

$$A_{ij} = \begin{cases} \dfrac{\partial^2\varphi}{\partial\mu_i\partial\mu_j}, \quad i,j=1,\cdots,M, \\ \dfrac{\partial^2\varphi}{\partial\mu_i\partial\lambda} = -1, \quad i=1,\cdots,M, j=M+1, \\ \dfrac{\partial^2\varphi}{\partial\lambda^2} = 0, \quad i=M+1, j=M+1, \end{cases} \tag{16.6.5}$$

而 $(M+1)\times N$ 阶矩阵 $\underset{\sim}{B}$ 为

$$B_{ij} = \begin{cases} \dfrac{\partial^2\varphi}{\partial\mu_i\partial n_j}, \quad i=1,\cdots,M, j=1,\cdots,N, \\ \dfrac{\partial^2\varphi}{\partial\lambda\partial n_j} = 1, \quad i=M+1, j=1,\cdots,N. \end{cases} \tag{16.6.6}$$

利用误差传递公式, 估计量 $\hat{\boldsymbol{\mu}}$ 的协方差矩阵 $U_{ij}=\mathrm{cov}(\hat{\mu}_i,\hat{\mu}_j)$ 可以通过数据 $\boldsymbol{n}$ 的协方差矩阵 $V_{ij}=\mathrm{cov}(n_i,n_j)$ 求得：

$$\mathrm{cov}(\hat{\mu}_i,\hat{\mu}_j) = \sum_{k,l=1}^{N}\frac{\partial\hat{\mu}_i}{\partial n_k}\frac{\partial\hat{\mu}_j}{\partial n_l}\mathrm{cov}(n_k,n_l). \tag{16.6.7}$$

式 (16.6.7) 中的偏导数可利用式 (16.6.4) 计算：

$$\frac{\partial\hat{\mu}_i}{\partial n_k} = -\left(\underset{\sim}{A}^{-1}\underset{\sim}{B}\right)_{ik} \equiv C_{ik}, \tag{16.6.8}$$

式中矩阵 $\underset{\sim}{A}$ 和 $\underset{\sim}{B}$ 由方程式 (16.6.5) 和式 (16.6.6) 给定. 这里我们要使用的并不是整个矩阵 $\underset{\sim}{C}$, 而是其 $M\times N$ 阶子矩阵, 即将代表拉格朗日乘子的 $i=M+1$ 行排除在外. 于是, 协方差矩阵 $\underset{\sim}{U}$ 的最终形式可以表示为更紧凑的形式：

$$\underset{\sim}{U} = \underset{\sim}{C}\,\underset{\sim}{V}\,\underset{\sim}{C}^{\mathrm{T}}. \tag{16.6.9}$$

这里 $\underset{\sim}{V}$ 是 $N\times N$ 阶矩阵, 而 $\underset{\sim}{U}$ 是 $M\times M$ 阶矩阵.

方程式 (16.6.5) 和式 (16.6.6) 中的偏导数依赖于式 (16.4.7) $\varphi(\boldsymbol{\mu},\lambda)$ 中正规函数和对数似然函数的形式的选择, 例如泊松, 高斯 ($\ln L=-\chi^2/2$) 函数等. 对于观

测数据视为独立泊松变量、协方差矩阵为 $V_{ij} = \delta_{ij}\nu_i$, 并利用基于极大熵的正规函数式 (16.5.14) 的情形, 我们有

$$\begin{aligned}\frac{\partial^2\varphi}{\partial\mu_i\partial\mu_j} =& -\alpha\sum_{k=1}^{N}R_{ki}R_{kj}\frac{n_k}{\nu_k^2}\\ &+\frac{1}{\mu_{\text{tot}}^2}\left[1-\frac{\delta_{ij}\mu_{\text{tot}}}{\mu_i}+\ln\left(\frac{\mu_i\mu_j}{\mu_{\text{tot}}^2}\right)+2S(\boldsymbol{\mu})\right]\end{aligned} \tag{16.6.10}$$

以及

$$\frac{\partial^2\varphi}{\partial\mu_i\partial n_j} = \frac{\alpha R_{ji}}{\nu_j}. \tag{16.6.11}$$

矩阵 $\underset{\sim}{A}$ 和 $\underset{\sim}{B}$ (从而矩阵 $\underset{\sim}{C}$) 通过实验中得到的 $\boldsymbol{\mu}$ 和 $\boldsymbol{\nu}$ 的估计值计算偏导数式 (16.6.10) 和式 (16.6.11) 确定 (参见式 (16.6.5), 式 (16.6.6) 和式 (16.6.8)). 表 16.1 汇总了泊松和高斯对数似然函数所对应的偏导数的表示式. 应当指出, 对于高斯对数似然函数 (即最小二乘法) 的情形, 所使用的量总是指 $\ln L = -\chi^2/2$, 而不是 χ^2 自身. Tikhonov 正规函数和基于极大熵的正规函数的偏导数已在 16.5.1 节和 16.5.2 节给定.

表 16.1　泊松和高斯随机变量的对数似然函数及其偏导数

	泊松	高斯 (最小二乘函数)
$\ln L$	$\sum_i (n_i\ln\nu_i - \nu_i)$	$-\frac{1}{2}\sum_{i,j}(\nu_i - n_i)\left(\underset{\sim}{V}^{-1}\right)_{ij}(\nu_j - n_j)$
$\frac{\partial \ln L}{\partial\mu_i}$	$\sum_j\left(\frac{n_j}{\nu_j}-1\right)R_{ji}$	$-\sum_{j,k}R_{ji}\left(V^{-1}\right)_{jk}(\nu_k - n_k)$
$\frac{\partial^2\ln L}{\partial\mu_i\partial\mu_j}$	$-\sum_k\frac{n_kR_{ki}R_{kj}}{\nu_k^2}$	$-\left(\underset{\sim}{R}^T\underset{\sim}{V}^{-1}\underset{\sim}{R}\right)_{ij}$
$\frac{\partial^2\ln L}{\partial n_i\partial\mu_j}$	$\frac{R_{ij}}{\nu_i}$	$\left(\underset{\sim}{V}^{-1}\underset{\sim}{R}\right)_{ij}$

为了确定偏差 $b_i = E(\hat{\mu}_i) - \mu_i$, 我们可以利用近似关系式 (11.6.4) 计算期望值 $E(\hat{\mu}_i)$:

$$b_i = E(\hat{\mu}_i) - \mu_i \approx \tilde{\mu}_i + \sum_{j=1}^{N}C_{ij}(\nu_j - \tilde{n}_j) - \mu_i. \tag{16.6.12}$$

将方程式 (11.6.4) 的 μ_i 的估计量代入, 并将 ν_j 用相应的估计量 $\hat{\nu}_j = \sum_k R_{jk}\hat{\mu}_k$ 代入, 求得

$$\hat{b}_i = \sum_{j=1}^{N}C_{ij}(\hat{\nu}_j - n_j) = \sum_{j=1}^{N}\frac{\partial\hat{\mu}_i}{\partial n_j}(\hat{\nu}_j - n_j). \tag{16.6.13}$$

计算 $\hat{b}_i$ 时用到的近似关系在小 $(\hat{\nu}_j - n_j)$ 情形下, 等价地在大的正规参数 α 的情形下是成立的. 对于小的正规参数 α, 矩阵 $\underset{\sim}{C}$ 事实上趋向于 0, 从而由式 (11.6.8) 可

知估计量 $\hat{\mu}_i$ 与测量值 n_j 无关. 然而这种情形下, 偏差实际上达到其最大值. 但是, 因为我们只是为了利用 $\hat{b}_i$ 及其方差来确定正规参数 α, 对于该目的而言, 这一近似已经足够.

利用误差传递公式 (忽略矩阵 $\underset{\sim}{C}$ 的方差), 我们得到 $\hat{b}_i$ 的协方差矩阵 $\underset{\sim}{W}$:

$$W_{ij} = \mathrm{cov}\left(\hat{b}_i, \hat{b}_j\right) = \sum_{k,l=1}^{N} C_{ik} C_{jl} \mathrm{cov}\left[\left(\hat{\nu}_k - n_k\right), \left(\hat{\nu}_l - n_l\right)\right]. \tag{16.6.14}$$

利用 $\hat{\nu}_k = \sum_m R_{km}\hat{\mu}_m$ 将协方差矩阵 $\mathrm{cov}\left(\hat{\nu}_k, \hat{\nu}_l\right)$ 与真值分布估计量的协方差矩阵 $U_{ij} = \mathrm{cov}\left(\hat{\mu}_i, \hat{\mu}_j\right)$ 联系起来, 后者通过方程式 (11.6.9) $\underset{\sim}{U} = \underset{\sim}{C}\,\underset{\sim}{V}\,\underset{\sim}{C}^T$ 与观测数据的协方差矩阵 $\underset{\sim}{V}$ 联系起来. 将所有这些关系归纳在一起, 得出

$$\begin{aligned}\underset{\sim}{W} &= \left(\underset{\sim}{C}\,\underset{\sim}{R}\,\underset{\sim}{C} - \underset{\sim}{C}\right)\underset{\sim}{V}\left(\underset{\sim}{C}\,\underset{\sim}{R}\,\underset{\sim}{C} - \underset{\sim}{C}\right)^T \\ &= \left(\underset{\sim}{C}\,\underset{\sim}{R} - \underset{\sim}{I}\right)\underset{\sim}{U}\left(\underset{\sim}{C}\,\underset{\sim}{R} - \underset{\sim}{I}\right)^T,\end{aligned} \tag{16.6.15}$$

其中 $\underset{\sim}{I}$ 是 $M \times M$ 阶单位矩阵. 方差 $V\left(\hat{b}_i\right) = W_{ii}$ 可以告知我们, 所估计的偏差 $\hat{b}_i$ 是否显著地偏离 0; 这一点可以用作确定正规参数的一种准则.

但是, 在讨论正规参数问题之前, 需要着重指出, 对于所有的正规化去弥散方法, 偏差 $\hat{b}_i$ 一般不为 0; 也就是说, 偏差一般是真值分布的某个函数, 并非处处为 0. 然而它的数值对于某些特定的 $\boldsymbol{\mu}$ 实际上可以是 0. 去弥散方法的一个指导性原则是, 如果真值 $\boldsymbol{\mu}$ 具有某种先验地认定是正确的分布性质, 则应选择一种使得偏差为 0(或很小) 的方法. 例如, 如果真值分布是均匀分布, 则基于 $k=1$ 的 Tikhonov 正规函数式 (16.5.2) 的估计将具有 0 偏差; 如果真值分布是线性分布, 则基于 $k=2$ 的 Tikhonov 正规函数式 (16.5.3) 的估计将具有 0 偏差; 如此等等. 如果真值分布等于参考分布 $\boldsymbol{q}$, 则利用式 (16.5.23) 交叉熵的去弥散方法将具有 0 偏差.

16.7　正规参数的选择

正规参数 α 的选择, 等价地 $\Delta \ln L$(或 $\Delta\chi^2$) 值的选择决定了估计量 $\hat{\boldsymbol{\mu}}$ 的偏差和方差之间的平衡. 当设定很大的 α 值时, 由式 (16.4.7) 可知, 似然函数对于去弥散问题的解起主要作用, 对 $\varphi\left(\boldsymbol{\mu}, \lambda\right)$ 求极大等价于对 $\ln L$ 求极大, 故有 $\ln L = \ln L_{\max}$(或者利用最小二乘法有 $\chi^2 = \chi^2_{\min}$) 的关系, 并且相应地估计量 $\hat{\boldsymbol{\mu}}$ 有非常大的方差. 在另一种极端情形下 $\alpha \to 0$, 正规函数 $S\left(\boldsymbol{\mu}\right)$ 对问题的解起支配作用, 导致完全平滑的解.

估计量 $\hat{\boldsymbol{\mu}}$ 的偏差和方差之间的最优平衡可以有不同的定义; 它们涉及的因素包括对于协方差矩阵 $U_{ij} = \mathrm{cov}\left(\hat{\mu}_i, \hat{\mu}_j\right)$ 和偏差 $\hat{b}_i$ 的估计, 以及偏差的估计量的协方差矩阵 $W_{ij} = \mathrm{cov}\left(\hat{b}_i, \hat{b}_j\right)$. 这里 $\underset{\sim}{U}$ 和 $\underset{\sim}{W}$ 指的是估计值 $\underset{\sim}{\hat{U}}$ 和 $\underset{\sim}{\hat{W}}$(字母上的 $\wedge$ 被省略).

去弥散估计量 $\hat{\boldsymbol{\mu}}$ 的优良程度可用对于所有子区间求平均的**均方误差** MSE 来表示：

$$\mathrm{MSE} = \frac{1}{M}\sum_{i=1}^{M}\left(U_{ii} + \hat{b}_i^2\right). \tag{16.7.1}$$

其中 $\sum_{i=1}^{M} U_{ii}/M$ 和 $\sum_{i=1}^{M} \hat{b}_i^2/M$ 分别称为**均方差**和**均方偏差**, 即均方误差等于均方差与均方偏差之和. 如果我们要求 α 的选择能使 MSE 等于某个特定的数值, 则 α 的确定需要通过数值计算来实现. 通常首先选择一个 α 的初值, 对 $\varphi\left(\boldsymbol{\mu}, \lambda\right)$ 求极大, 进行迭代计算, 直到找出满足要求的解.

但是人们可能会认为, 不同的子区间对于 MSE 的贡献大小应当按每个子区间中的测量精度来确定. 因为均值 μ_i 的泊松变量的方差等于 μ_i, 与最小二乘法中的 χ^2 类似, 我们可以定义一个**加权均方误差** MSE′：

$$\mathrm{MSE}' = \frac{1}{M}\sum_{i=1}^{M}\frac{U_{ii} + \hat{b}_i^2}{\hat{\mu}_i}. \tag{16.7.2}$$

对于服从泊松分布的测量数据, 量 MSE′ 代表了由于有限分辨导致的均方误差, 因此要求这个量是一个小量是合理的.

正规参数的一种常用的选择方法基于如下考虑：平均地每个子区间对于 χ^2 的贡献应当近似等于 1, 也就是说, α 值的选择应当使得 $\chi^2 = N$. 这一要求推广到对数似然函数的情形即是 α 值的选择应当使得 $\Delta \ln L = \ln L_{\max} - \ln L = N/2$, 因为对于服从高斯分布的测量数据 $\boldsymbol{n}$, 关系式 $\ln L = -\chi^2/2$ 成立.

一种朴素的想法是, 式 (16.4.2) 中的 $\Delta\chi^2$ 选择为等于 1 表示数据 $\boldsymbol{n}$ 与其估计值 $\hat{\boldsymbol{\mu}}$ 之间的差别 (或一致性) 达到一种比较适当的程度. 但是这种做法求得的问题的解其方差大到不合理的程度. 其原因是：估计量 $\hat{\nu}_i$ 不仅接收到第 i 子区间的贡献 n_i, 而且同时还接收到其相邻各子区间的贡献. 估计量 $\hat{\nu}_i$ 与测量值 n_j 之间的耦合可表示为

$$\frac{\partial \hat{\nu}_i}{\partial n_j} = \frac{\partial}{\partial n_j}\sum_{k=1}^{M} R_{ik}\hat{\mu}_k = \left(\underset{\sim}{R}\,\underset{\sim}{C}\right)_{ij}. \tag{16.7.3}$$

基于上述思想, 文献 [153] 提出了对于准则 $\Delta\chi^2 = 1$ 的一种修正. 它要求等效 χ^2 量的数值增加 1：

$$\Delta\chi^2_{\mathrm{eff}} = \left(\hat{\boldsymbol{\nu}} - \boldsymbol{n}\right)^T \underset{\sim}{R}\,\underset{\sim}{C}\,\underset{\sim}{V}^{-1}\left(\underset{\sim}{R}\,\underset{\sim}{C}\right)^T\left(\hat{\boldsymbol{\nu}} - \boldsymbol{n}\right) = 1, \tag{16.7.4}$$

式中矩阵 $\underset{\sim}{R}\underset{\sim}{C}$ 有效地降低了估计量 $\hat{\nu}_i$ 与数据 n_i 之间的耦合强度.

另一种途径是我们可以检查估计量 $\hat{\boldsymbol{\mu}}$ 的偏差及其方差的估计值. 如果偏差显著地偏离 0, 那么我们可以将这一偏差加以扣除, 这等价于减小 $\Delta\ln L$ 值. 我们可以利用各子区间偏差平方的加权和作为偏差大小的一种度量:

$$\chi_b^2 = \sum_{i=1}^{M} \frac{\hat{b}_i^2}{W_{ii}}. \tag{16.7.5}$$

求解问题的准则是减小 $\Delta\ln L$(即增大 α) 直到 χ_b^2 等于一个充分小的值, 例如与子区间数 M 相等. 当 $\chi_b^2 = M$ 时, 偏差的标准差与偏差值本身近似相等, 因此进一步减小偏差将不能对误差有所改善.

图 16.2 显示了 $\Delta\ln L$ 与估计量 $\hat{\boldsymbol{\mu}}$ 的均方差 (mean variance)、均方偏差 (mean squared bias) 以及两者之和即均方误差 (mean squared error) 之间的函数关系. 它们对应于图 16.1 所示的例子. 图 16.1 中去弥散通过响应矩阵的求逆实现, 而图 16.2 中去弥散通过 (a) 极大熵原理 (b)Tikhonov 正规函数实现. 小 $\Delta\ln L$ 值处均方偏差估计值的增大反映了估计量 $\hat{b}_i$ 方差的增大; 当 $\Delta\ln L$ 逼近 0 时真实的偏差亦趋近 0. 图中的箭头指示出基于上面讨论的不同准则求得的解的位置; 下一节中将对它们做进一步的讨论.

根据奇异值分析 [143] (singular value analysis) 或利用交叉验证 [158] (cross-validation) 的步骤, 提出了选择正规参数的其他一些准则. 令人遗憾的是, 时至今日, 正规参数 α 的选择仍然是一个有待解决的问题. 事实上, 如图 16.2 所示, 当 $\Delta\ln L$ 逐渐减小时, 均方误差亦平稳地减小; 直到某一点达到极小, 进一步减小 $\Delta\ln L$ 会使均方误差迅速增大. 极小点对应的 α 值应该是 α 的自然选择.

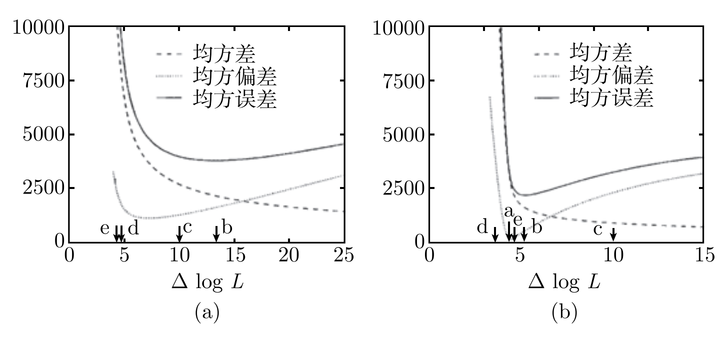


图 16.2 估计量 $\hat{\boldsymbol{\mu}}$ 的均方差、均方偏差、均方误差的估计值与 $\Delta\ln L$ 的函数关系: (a) 极大熵原理; (b) Tikhonov 正规函数 ($k = 2$)

箭头指示图 16.3 和图 16.4 中的解的位置: ‘b’ 是最小 MSE 的解, ‘c’ 是 $\Delta\ln L = N/2$ 的解, ‘d’ 是 $\Delta\chi^2_{\text{eff}} = 1$ 的解, ‘e’ 是 $\chi_b^2 = M$ 的解. 对于极大熵原理的情形, $\Delta\ln L = 970$ 的 Bayes 解在图中没有画出. 对于 Tikhonov 正规函数的情形, 箭头 ‘a’ 指示加权 MSE 极小值对应的解

16.8 去弥散计算实例

图 16.3 和图 16.4 分别显示了利用极大熵原理和 Tikhonov 正规函数求解去弥散问题的实例. $\boldsymbol{\mu}, \boldsymbol{\nu}$ 和 $\boldsymbol{n}$ 的分布与图 16.1(a)~(c) 中所示相同, 子区间数 $N = M=20$, 所有的效率皆等于 1, 本底 $\boldsymbol{\beta}$ 等于 0. 对式 (16.4.7) 定义的函数 $\varphi(\boldsymbol{\mu}, \lambda)$ 求极大找到估计量 $\hat{\boldsymbol{\mu}}$, 这里 $\varphi(\boldsymbol{\mu}, \lambda)$ 是将各子区间的观测数据视为独立的泊松变量、利用相应的对数似然函数构建而成的. 原始的 “真值” 直方图 $\boldsymbol{\mu}$ 与去弥散之后的解 $\hat{\mu}_i$ 及其误差 $\sqrt{U_{ii}}$(对应于给定值的正规参数 α 或给定值的 $\Delta \ln L$) 一起画在图的左边. 图的右边则是偏差 b_i 及其标准离差 $\sqrt{W_{ii}}$ 的相应估计值. 后者不应与真实残差 $\hat{\mu}_i - \mu_i$ 相混淆, 如果没有真值直方图 $\boldsymbol{\mu}$ 的知识, 真实残差是不可能知晓的. 而偏差的估计 $\hat{b}_i$ 则是根据观测数据确定的.

首先考察图 16.3, 应用极大熵原理的正规函数 $S(\boldsymbol{\mu}) = H(\boldsymbol{\mu})$ 求解去弥散问题. 图 16.3(a) 对应于 $\alpha = 1/\mu_{\text{tot}}$, 即式 (16.5.19) 所示的贝叶斯方案, 相应于 $\Delta \ln L = 970$. 我们选择以此为例是为了说明先验密度 $\pi(\boldsymbol{\mu}) = \Omega(\boldsymbol{\mu})$ 不能得到合理的解. 尽

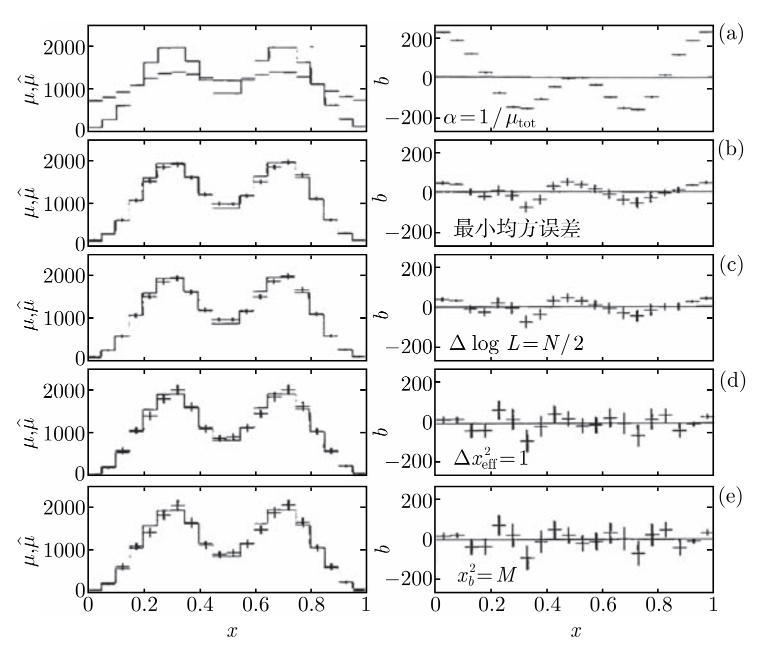


图 16.3 左图显示真值分布 (直方图) 和选择不同的正规参数 α 值利用极大熵去弥散求得的分布 (带误差杆的点); 右图显示相应的偏差估计值

其中 (a) 为贝叶斯方案 $\alpha = 1/\mu_{\text{tot}}$ 的解; (b) 最小均方误差解; (c) $\Delta \ln L = N/2$ 解; (d) $\Delta\chi^2_{\text{eff}} = 1$ 解; (e) $\chi^2_b = M$ 解. 本例中, 最小加权均方误差解与 (c) $\Delta \ln L = N/2$ 解类似

管标准差 $\sqrt{U_{ii}}$ 很小, 但估计值 $\hat{\mu}_i$ 偏差很大. 图右边显示的偏差估计值 $\hat{b}_i$ 确实很大, 与误差杆 $\sqrt{W_{ii}}$ 比较我们可以看到, $\hat{b}_i$ 明显地偏离于 0. 应当指出, 这里所估计的偏差实际上与真实残差 $\hat{\mu}_i - \mu_i$ 并不很好地一致, 这是由于在构建估计量 $\hat{b}_i$ 时引入了近似之故, 参见方程式 (16.6.4) 和式 (16.6.13). 当 $\Delta \ln L$ 变小时近似变好, 直到偏差估计值 $\hat{b}_i$ 的标准差 $\sqrt{W_{ii}}$ 变得与偏差值相当, 这种关系都成立.

图 16.3(b) 显示了式 (16.7.1) 表示的均方误差 (MSE) 达到极小所对应的结果. 这对应于 $\Delta \ln L = 13.3$ 以及 $\chi_b^2 = 154$. 虽然估计的偏差大大小于 $\alpha = 1/\mu_{\text{tot}}$ 得到的结果, 但仍然明显地偏离于 0.

对应于 $\Delta \ln L = N/2 = 10$ 的解示于图 16.3(c), 相应的 χ_b^2 值为 87. 这种情形下的偏差比最小 MSE 的结果要小一些, 但仍然明显地偏离于 0. 在这一特定的例子中利用式 (16.7.2) 的最小加权均方误差给出 $\Delta \ln L = 10.5$, 因此与 $\Delta \ln L = N/2 = 10$ 的结果类似.

对应于 $\Delta\chi^2_{\text{eff}} = 1$ 和 $\chi_b^2 = M$ 的解分别示于图 16.3(d) 和 16.3(e). 这两种情形下偏差接近于 0, 但是其代价是 $\hat{\mu}_i$ 的方差和 $\hat{b}_i$ 的方差比 $\Delta \ln L = N/2$ 或最小 MSE 的结果要大. $\Delta\chi^2_{\text{eff}} = 1$ 情形下 $\chi_b^2 = 20.8$, 而 $\chi_b^2 = M$ 情形下 $\Delta\chi^2_{\text{eff}} = 0.85$, 所以在这一例子中, 两者实际上非常类似.

现在考察图 16.4, 其中的真值分布与图 16.3 相同, 但现在利用 $k = 2$ 的 Tikhonov 正规函数求解去弥散问题. 该图对应于 (a) 最小加权均方误差解, (b) 最小均方误差解, (c) $\Delta \ln L = N/2$ 解, (d) $\Delta\chi^2_{\text{eff}} = 1$ 解, (e) $\chi_b^2 = M$ 解. 本例中方法 (a) 、(b) 和 (e) 得到了较好的结果, 去弥散后的估计值 $\hat{\boldsymbol{\mu}}$ 与真值 $\boldsymbol{\mu}$ 接近, 且 $\hat{\boldsymbol{\mu}}$ 的统计误差 $\sqrt{U_{ii}}$ 小, 偏差 $\hat{b}_i$ 接近于 0 且其统计误差 $\sqrt{W_{ii}}$ 合理地小. 方法 (c)$\Delta \ln L = N/2$ 的解尽管统计误差 $\sqrt{U_{ii}}$ 相当小, 但偏差很大且显著地偏离于 $0(b_i^2 \gg W_{ii})$. 但方案 (d) 看起来离我们的要求过远：虽然其偏差接近于 0, 但并不比 $\chi_b^2 = M$ 解的偏差更接近于 0; 而统计误差则比 $\chi_b^2 = M$ 解大得多.

从图 16.4 左半图的最后一个子区间可以发现与 Tikhonov 正规函数相关的一个问题：除了图 16.4(a) 之外, 最后一个子区间的估计值都是负值. 事实上, 该算法不能保证不出现负值. 如果必须不出现负值, 那么算法必须加以修改, 比如人为地减小可能出现负估计值的那些点的误差. 在基于极大熵原理的去弥散方法中不存在这样的问题, 因为当任一 μ_i 趋近 0 时 $S(\boldsymbol{\mu})$ 的梯度发散. 这种罚函数的行为保证了 μ_i 的正值性.

本章讨论的技巧容易推广到多维分布的情形. 例如对于二维的情形, 去弥散方法已广泛地应用于影像复原 [140,141,142,154], 以及天文学 [148] 和医学 [147] 中的成像问题. 对于此问题的详尽讨论超出了本书的范围, 我们只用一个简单的例子来说明一些基本概念.

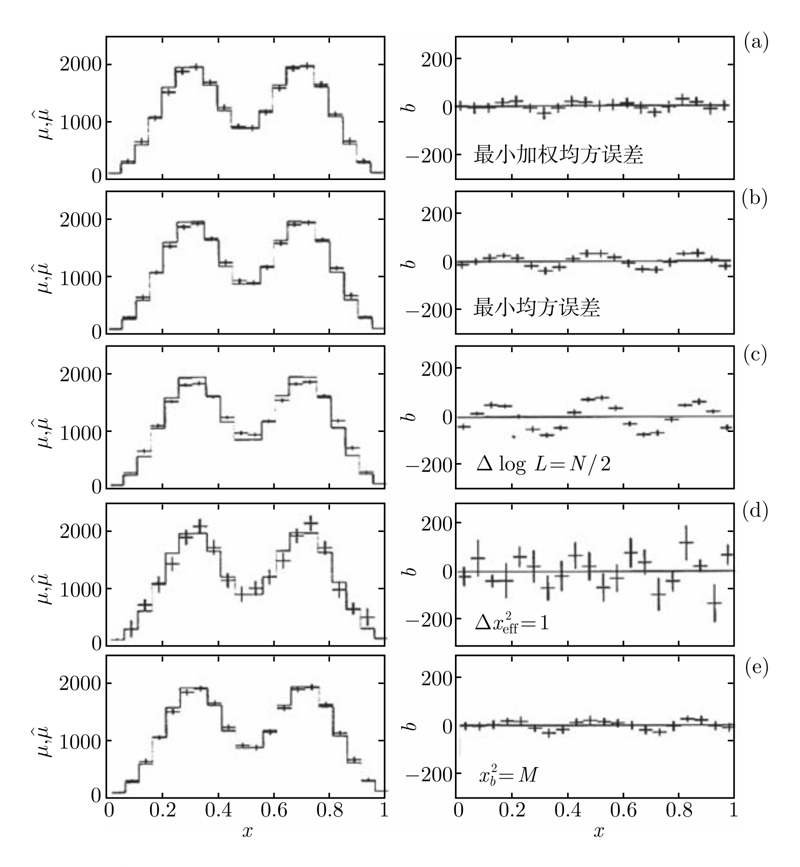


图 16.4 左图显示真值分布 (直方图) 和选择不同的正规参数 α 值、利用 Tikhonov 正规函数 (k=2) 去弥散求得的分布 (带误差杆的点); 右图显示相应的偏差估计值

其中 (a) 为最小加权均方误差解; (b) 最小均方误差解; (c) $\Delta \ln L = N/2$ 解; (d) $\Delta\chi^2_{\text{eff}} = 1$ 解; (e) $\chi^2_b = M$ 解

图 16.5 显示了一个基于极大熵原理的去弥散方法应用于一张 56×56 个像素点的照片的实例. 图 16.5(a) 是其 "真实" 影像, 用向量 $\boldsymbol{\mu}$ 表示. 图 16.5(b) 是 "真实" 影像利用标准差等于 0.6 倍像素大小的高斯分辨函数加以模糊化后的影像.

假设, 照片中最亮的像素内包含的有效 "光子" 数 (或卤化银颗粒数、光电子数等, 取决于成像系统的种类) 设置为 10^4 量级. 如果像素 i 中的计数视为期望值 ν_i 的泊松变量 n_i, 则照片的亮区处计数的相对统计涨落具有 1%的量级 ($\sigma_i/\nu_i = 1/\sqrt{\nu_i}$). 图 16.5(c) 显示了基于极大熵原理去弥散后的影像, 其中 $\Delta \ln L = N/2$, $N = 3136$ 是像素总数. 图 16.5 右侧的直方图表示各张照片第 36 行像素的光强度分布.

对于这一特定的例子, 极大熵方法较之 Tikhonov 正规函数具有某种优越性. 首先, 极大熵去弥散具有前面提到过的特点, 即去弥散之后所有的子区间都有正

值. 其次, 该方法的优点是熵可以直接推广到多维分布的情形, 这是因为熵 $H = -\sum_j p_j \ln p_j$ 等于各子区 $p_j \ln p_j$ 值的简单求和. 这一事实所导致的结果是, 极大熵去弥散方法对于出现孤立的峰的惩罚程度是相对轻微的. 大的峰值在影像中呈现为星状的亮点, 这是极大熵方法被天文学家广泛采用的原因之一.

极大熵方法的一个缺点是, 它必须求解关于 $\boldsymbol{\mu}$ 的一组非线性方程. 但是一张照片中的像素数对于通过矩阵直接求逆求解而言一般是过于大了, 因此去弥散最终必定需要求助于迭代数值计算.

而对 Tikhonov 正规函数方法, 我们可以利用拉普拉斯算子的有限差分近似将函数 $S(\boldsymbol{\mu})$ 推广到二维的情形, 例如参考文献 [150] 的 18 章. 对于图 16.3 和 16.4 所示的相对平滑的分布, 对于一给定的偏差, Tikhonov 正规函数方法给出较小的方差. 而对具有尖峰的分布, 如图 16.5 所示的照片那样, 方差就不那么小了.

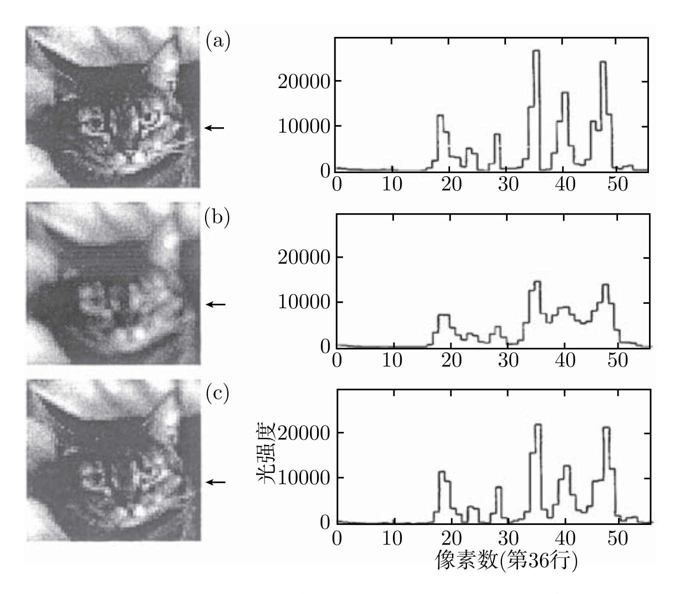


图 16.5　(a) 原始 “真实” 影像 $\boldsymbol{\mu}$; (b) “真实” 影像 $\boldsymbol{\mu}$ 利用标准差等于 0.6 倍像素大小的高斯分辨函数加以模糊化后得到的观测影像 $\boldsymbol{n}$; (c) 利用极大熵原理去弥散后的影像

右侧的直方图表示各张照片第 36 行 (如箭头所指) 各像素的光强度分布

16.9　数 值 计 算

前面各节描述的去弥散方法的数值计算可能是一个不太容易的任务. 找到函数

$$\varphi(\boldsymbol{\mu}, \lambda) = \alpha \ln L(\boldsymbol{\mu}) + S(\boldsymbol{\mu}) + \lambda \left[n_{\text{tot}} - \sum_{j=1}^{N} \nu_j \right] \tag{16.9.1}$$

关于 $\boldsymbol{\mu}$ 和拉普拉斯乘子 λ 的极小值需要求解式 (16.6.1) 所示的 $M+1$ 个方程. 如果 φ 是 $\boldsymbol{\mu}$ 的二次函数, 则式 (16.6.1) 为线性方程组. 例如, 当观测数据 $\boldsymbol{n}$ 为高斯分布, 则我们有对数似然函数 $\ln L = -\chi^2/2$, 再运用 Tikhonov 正规函数, 就属于这种情形. 文献 [143] 讨论了关于这类问题的基于奇异值分解 (singular value decomposition) 的求解方法. 如果 φ 包含泊松分布的对数似然函数, 或包含基于熵极大的正规函数, 则式 (16.6.1) 为非线性方程组, 必须通过迭代数值计算求解.

作为例子, 我们考察泊松分布的对数似然函数的情形, 按照式 (16.2.9) 和式 (16.2.10) 可得

$$\ln L(\boldsymbol{\mu}) = \sum_{i=1}^{N} (n_i \ln \nu_i - \nu_i), \tag{16.9.2}$$

正规函数取为 $S = H$, H 为式 (16.5.11) 所示的熵.

图 16.6 显示了一种可能的极大熵正规函数的解法. 图中三根轴表示 $\boldsymbol{\mu}$ 空间中的三维, 斜的平面 (图中用三角形表示) 是满足 $\sum_i \nu_i = n_{\text{tot}}$ 关系的子空间. 该平面中箭头指示的两个点是极大熵点 (所有 μ_i 相等) 和极大似然点. 这两点间的连接曲线代表方程式 (16.4.7) 所有可能的解, 这些解对应于不同的正规参数 α 值; 例如, $\alpha = 0$ 对应于极大熵点, $\alpha \to \infty$ 则对应于极大似然点. 该曲线通过常数熵等值线 (图中以极大熵点为中心的闭合曲线) 与常数似然值等值线 (以极大似然点为中心的闭合曲线) 的交点. 应当指出, 极大似然点并不处于所有 $\mu_i > 0$ 的容许区. 实际上, 这是典型的振荡型极大似然解, 参见图 16.1(d).

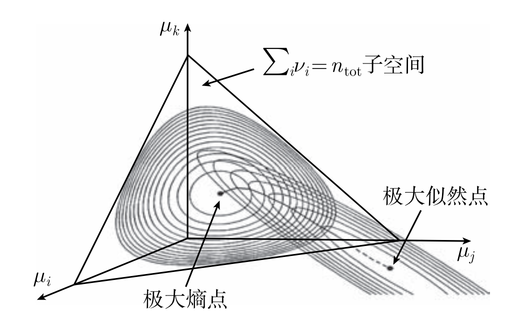


图 16.6 极大熵去弥散数值计算的三维 $\boldsymbol{\mu}$ 空间图示

一般而言极大似然点不能用作 $\boldsymbol{\mu}$ 的初值, 因为在那一点 μ_i 通常为负值, 因此熵没有定义. 极大熵点则可取为初值, 这时要求所有 μ_i 相等, 并满足约束

$$\begin{aligned}\nu_{\text{tot}} &= \sum_{i=1}^{N} \nu_i = \sum_{i=1}^{N} \sum_{j=1}^{M} R_{ij} \mu_j = \sum_{j=1}^{M} \varepsilon_j \mu_j \\ &= n_{\text{tot}},\end{aligned} \tag{16.9.3}$$

式中 ε_j 是子区间 j 中的效率. 因此极大熵点为

$$\mu_i = \frac{n_{\mathrm{tot}}}{\sum_{j=1}^{M} \varepsilon_j}. \tag{16.9.4}$$

如果利用关系式 $S(\boldsymbol{\mu}) = \mu_{\mathrm{tot}} H$, 且各子区间效率 ε_j 不全等, 则 $S(\boldsymbol{\mu})$ 极大的 $\boldsymbol{\mu}$ 分布不是均匀分布, 而是由 M 个方程的解给定：

$$\ln \frac{\mu_i}{\mu_{\mathrm{tot}}} + \frac{S(\boldsymbol{\mu})\, \varepsilon_i}{n_{\mathrm{tot}}} = 0, \quad i = 1, \cdots, M \tag{16.9.5}$$

从 $S(\boldsymbol{\mu})$ 极大值点开始, 在 ν_{tot} 等于常数的子空间中沿着 φ 极大值的曲线搜索. 在此子空间中, 只需要对量

$$\Phi(\boldsymbol{\mu}) = \alpha \ln L(\boldsymbol{\mu}) + S(\boldsymbol{\mu}) \tag{16.9.6}$$

求极大即可, $\Phi(\boldsymbol{\mu})$ 等于式 (16.4.7) 中的 $\varphi(\boldsymbol{\mu})$ 去除含拉格朗日乘子 λ 的项. 在这一方法中, 拉格朗日乘子不会显式地出现在算法中. 但是直接要求 $\nabla\Phi = 0$ 将不能得到所要求的解. 而是首先需要将 $\nabla\Phi$ 投影到 ν_{tot} 等于常数的子空间, 然后要求该投影向量的各分量等于 0; 也即要求

$$D\Phi = \nabla\Phi - \boldsymbol{u}(\boldsymbol{u} \cdot \nabla\Phi) = 0 \tag{16.9.7}$$

可求得问题的解. 这里微分算子 D 表示梯度在 ν_{tot} 等于常数的子空间中的投影, $\boldsymbol{u}$ 是 $\nabla\nu_{\mathrm{tot}}$ 方向上的单位向量. 由式 (16.9.3) 立即有：

$$\frac{\partial \nu_{\mathrm{tot}}}{\partial \mu_k} = \sum_{i=1}^{N} \sum_{j=1}^{M} R_{ij} \frac{\partial \mu_j}{\partial \mu_k} = \varepsilon_k, \tag{16.9.8}$$

这样, 向量 $\boldsymbol{u}$ 即是归一化为单位长度的效率向量：

$$\boldsymbol{u} = \frac{\boldsymbol{\varepsilon}}{|\varepsilon|}. \tag{16.9.9}$$

于是, 我们从极大熵点开始, 并沿着 $D \ln L$ 方向前进一小步. 计算所得的 $\boldsymbol{\mu}$ 一般不直接落在 Φ 极大值的曲线上, 但只要步长充分小, $\boldsymbol{\mu}$ 将靠近 Φ 极大值的曲线. 我们可以检查 $|D\Phi|$ 的值作为计算所得的 $\boldsymbol{\mu}$ 与该曲线距离的度量. 如果 $|D\Phi|$ 超出某个给定的限, 那表示步长过大; 需要采用更小的步长重新搜索.

如果计算所得的 $\boldsymbol{\mu}$ 落在 Φ 极大值的曲线上, 那么关系 $\alpha D \ln L + DS = 0$ 成立, 对应的正规参数则为

$$\alpha = \frac{|DS|}{|D \ln L|}. \tag{16.9.10}$$

参数 α 可以直接设置为式 (16.9.10) 右边的数值, 然后向曲线 $D\Phi = 0$ 进行搜索. 这一步骤可利用函数求极大的标准方法 (通常称为极小化, 参见文献 [38], [150]) 来完成. 这一搜索步骤同样是在子空间 $\nu_{\rm tot} = n_{\rm tot}$ 中进行的, 即搜索方向是在该子空间中的投影方向. 搜索的方式是, 沿着 $D \ln L$ 方向前进一个步长, 相应地增加了 α 值, 向解 $D\Phi = 0$ 逼近, 直至达到所要求的 $\Delta \ln L = \ln L_{\max} - \ln L$ 值为止. 计算的中间结果可以存储起来并进行检查, 以确定计算的最优停止点.

虽然上述算法的基本概念也可应用于 Tikhonov 正规化方法, 但由于 $S(\boldsymbol{\mu})$ 极大的解不是唯一确定的, 在那里情况变得较为复杂. 例如, 对于 $k = 2$ 的情形, 任意线性函数都给出 $S = 0$. 我们可以直接从 $\mu_i = n_{\rm tot}/M$ 这一点开始搜索, 设定一个充分大的正规参数 α 值, 就能找到唯一解.

利用 Tikhonov 正规化方法而不要求满足条件 $\sum_i \nu_i = n_{\rm tot}$ 也是可能的, 因为在这里正规函数并不倾向于将问题的解引向与 $n_{\rm tot}$ 很不同的归一化常数. 但是当我们略去归一化条件后, 将得不到严格的 $\sum_i \hat{\nu}_i = n_{\rm tot}$ 关系. 我们可以说, $\hat{\nu}_{\rm tot}$ 应当是事例总数 $n_{\rm tot}$ 的一个无偏估计量, 但因为 $\hat{\nu}_{\rm tot}$ 一般与 $n_{\rm tot}$ 的偏差不大, 这一归一化条件的约束通常可以不做要求.

参 考 文 献

概率和统计

[1] Kendal M, Stuart A. The Advanced Theory of Statistics. London: Charles Griffin & Company Limited, 1963, 1967, 1966, 1, 2, 3

[2] Cramer H. Mathematical Methods of Statistics. Princeton: Princeton University Press, 1966

[3] Fisher R. Statistical Methods for Research Workers. Edinburgh-London: Oliver and Boyd, 1958

[4] Feller W. An Introduction to Probability Theory and Its Applications. John Wiley and Sons, 1966, 1968. vol. 1.2

[5] Bradley J. Distribution-Free Statistical Tests. New Jersey: Englewood Cliffs, Prentice-Hall, Inc., 1968

[6] Breiman L. Statistics With a View Toward Applications. Boston: Houghton Mifflin Company, 1973

[7] Lindley D. Introduction to Probability and Statistics. London: Cambridge University Press, 1965. vol. 1.2

[8] Mood A, Grayball F, Boes D. Introduction to the Theory of Statistics. New York: McGraw-Hill Book Company, Inc., 1974

[9] Walpole R. Introduction to Statistics. New York: Macmillan Publishing Co., Inc., London: Collier Macmillan Publishers, 1982

[10] Hogg R, Craig A. Introduction to Mathematical Statistics. New York: Macmillan Publishing Co., Inc., 1978

[11] Larsen R, Marx M. An Introduction to Mathematical Statistics and Its Application. Englewood Cliffs, New Jersey: Prentice-Hall, 1981

[12] Larson H. Introduction to Probability Theory and Statistical Inference. New York: John Wiley & Sons, 1982

[13] Neter J et al. Applied Statistics. Boston: Allyn and Bacon, Inc., 1982

[14] Rohatgi V. An Introduction to Probability Theory and Mathematical Statistics. New York: John Wiley & Sons, 1976

[15] Sachs L. Applied Statistics: A Handbook of Techniques. 2nd edition, New York: Springer-Verlag, 1982

[16] Rose C, Smith M D. Mathematical Statistics With Mathematica. Berlin: Springer, 2002

[17] Lehman E. Nonparametrics: Statistical Methods Based on Ranks. San Francisco: Holden Day, Inc., 1975

[18] Kruskal W, Wallis W. The use of ranks in one-criterion variance analysis. J. Amer. Statist. Ass., 1952, 47: 583

[19] David F. A χ^2 “smooth” test for goodness-of-fit. Biome-trika, 1947, 34: 299

[20] Wilcoxon F. Individual comparisions by ranking methods. Biometrics Bulletin, 1945, 1: 80

[21] Durbin J. Kolmogorov-Smilnov tests when parameters are estimated with applications to tests of exponentiality and tests on spacings, Biometrika, 1975, 62: 5

[22] Fisher R A. On the “probable error” of a coefficient of correlation deduced from a small sample. Metron, 1921 (1): Part4, 3–32

[23] Daniel W W. Biostatistics: A Foundation for Analysis in the Health Science. 7th New York: John Wiley & Sons. Inc., edition, 1999

[24] Spearman C. American Journal of Psychology, 1904, 15: 72–101

[25] Sheskin D J. Handbook of Parametric and Nonparametric Statistical Procedures. 3rd edition. Boca Raton, Chapman & Hall/CRC, A CRC Press Company, 2004

[26] Kandall M G. Biometrika, 1938, 30: 81–93

[27] Prokhorov A V. Kendall Coefficient of Rank Correlation//Hazewinkel. Michiel. Encyclopaedia of Mathematics. Springer, 2001

[28] Lindeman R H, et al., Intruduction to bivariate and multivariate analysis. Glenview, IL : Scott, Foresman and Company, 1980

[29] Kandall M G, Babington-Smith B. Annals of Mathematical Statistics, 1939, 10: 275–287

[30] 中山大学. 概率论与数理统计. 北京：人民教育出版社, 1980

[31] 王福保. 概率论及数理统计. 上海：同济大学出版社, 1984

[32] 茆诗松等. 高等数理统计. 北京：高等教育出版社, 1998

[33] 范金城, 吴可法编著. 统计推断导引. 北京：科学出版社, 2001

[34] 朱永生. 实验物理中的概率与统计. 第二版. 北京：科学出版社, 2006

概率和统计在物理学中的应用

[35] Eadie W, et al. Statistical Methods in Experimental Physics. Amsterdam-London: North-Holland Publishing Company, 1971. James F. Statistical methods in experimental physics. (2nd edition). World Scientific, 2006

[36] Cooper B. Statistics for Experimentalists. Oxford: Pergamin Press, 1969

[37] Bevington P. Data Reduction and Error Analysis for the Physical Sciences. New York: Mc Graw-Hill Book Company, 1974

[38] Brandt S. Statistical and Computational Methods in Data Analysis. 3rd edition. New York: Springer, 1999

[39] Frodesen A, et al. Probability and Statistics in Particle Physics. Universitetsforlaget, Bergen-Oslo-Troms.ϕ, 1979

[40] Martin B. Statistics for physicists. London-New York: Academic Press, 1971

[41] Dowdy S, Weardon S, Chilko D. Statistics for Research. 3th edition, New Jersey: John Wiley & Sons, Inc., 2004

[42] Mickey R M, Dunn O J, Clark V A. Applied Statistics, 3rd edition, New Jersey: John Wiley & Sons, Inc., 2004

[43] Solmits F. Analysis of experiments in particle physics. Ann. Rev. Nucl. Sci., 1964, 14: 375

[44] Green J, Margerison D. Statistical Treatment of Experimental Data. Amsterdam-Oxford-New York: Elsevier Scientific Publishing Company, 1978

[45] Bethea R. et al. Statistical Methods for Engineers and Scientists. New York and Basel: Marcel Dekker, Inc., 1975

[46] Box E. Statistics for Experimenters. New York: John Wiley & Sons, 1978

[47] Ogawa J. Statistical Theory of the Analysis of Experimental Designs. New York: Macel Dekker, Inc., 1974

[48] Mandel J. The Statistical Analysis of Experimental Data. New York: John Wiley & Sons, 1964

[49] Walpole R, Myers R. Probability and Statistics for Engineer and Scientists. New York: Macmillan Publishing Co., Inc., 1978

[50] Meyer S. Data Analysis for Scientists and Engineers. New York: John Wiley & Sons, Inc., 1975

[51] Björck A. Least squares methods in physics and engineering. CERN 81-16, Geneva, 1981

[52] Wind H. Function parametrization. CERN 72-21, Geneva, 1972

[53] James F. Determining the statistical significance of experimental results. CERN DD 81-02, Geneva, 1981

[54] Orear J. Notes on statistics for physics(revised). CLNS 82/511, Cornell Univ., Ithaca, 1982

[55] Hudson D. Lectures on elementary statistics and probability. CERN 63-29, Geneva, 1963; Statistic Lectures 2: Maximun Likehood and Least Square Theory, CERN 64-18, Geneva, 1964

[56] Lyons L. Statistics for Nuclear and Particle Physicists. Cambridge: Cambridge University Press, 1986

[57] Cowan G. Statistical Data Analysis. New York: Oxford University Press Inc., 1998

[58] Brandt S. Data Analysis. 3rd edition, Springer-Verlag New York Inc., 1999

[59] CERN 2000-005(2000)
[60] Proceedings of the Conference“Advanced Statistical Techniques in Particle Physics”, Grey College, Durham, UK, 2002
[61] 李惕碚. 实验的数学处理. 北京：科学出版社, 1983
[62] 赵特伟. 试验数据的整理与分析. 北京：铁道出版社, 1981
[63] Barlow R. arXiv Physics/0406120, 2004
[64] Nakamura K., et al. J. Phys. 2010, (G37): 075021
[65] Schmelling M. Phys. Scripta, 1995, 51: 676
[66] Ablikim M, et al. BES collaboration. Phys. Rev., 2004, (D70): 112007
[67] Neyman J. Phil. Trans. Royal Soc. London, Series, 1937(A236): 333; A Selection of Early Statistical Papers on J. Neyman. Berkeley: University of California Press, 1967: 250
[68] Feldman G, Cousins R. Phys. Rev., 1998(D57): 3873
[69] Roe B, Woodroofe M. Phys. Rev., 1999(D60): 053009
[70] Conrad J, et al. Phys. Rev., 2003(D67): 012002
[71] Bityukov S I, et al., Nucl. Instr. Meth. in Phys. Resear., 2000(A452): 518; Proc. of Conf. “Advanced statistical techniques in particle physics”, Durham, UK, 2002: 77
[72] I. Narsky. Nucl. Instr. Meth, 2000(A450): 444
[73] 朱永生. 高能物理与核物理, 2006, 30: 331
[74] CMS Collaboration. CMS physics technical design report, volume II: Physics performance, J. Phys., 2007 (G34): 995-1579, Appendix A.1
[75] 高原宁. Combine Measurements of α_S-How to Average Correlated Data. //Shan Jin and Yong-Sheng Zhu, Error Treatment in Particle Physics Experiments. CCAST –WL Workshop series: Vol 160, 2004
[76] 朱永生. 实验数据多元统计分析. 北京：科学出版社, 2009
[77] Frühwirth. 高能物理数据分析. 朱永生, 刘振安译. 合肥：中国科学技术大学出版社, 2011

贝叶斯统计

[78] Bayes T. Phil. Trans. Roy. Soc. 1763 (53):370
[79] Jeffreys H. Theory of Probability, 3rd edition, London: Oxford Univ. Press, 1961
[80] Savage L J. The Foundation of Statistics, 2nd revised edition, New York: Dover, 1972
[81] de Finetti B. Theory of Probability: A Critical Introductory Treatment (2 volums), New York: John Wiley, 1974
[82] Cox R T. Probability, frequency and reasonable expectation. Am. J. Phys., 1946(14):1
[83] Lindley D V. Introduction to Probability and Statistics from a Bayesian Viewpoint. Part 1: Probability, Part 2: Inference. Cambridge: Cambridge Univ. Press, 1965
[84] Lee P M. Bayesian Statistics: an Introduction. London: Edward Arnold, 1989

[85] Sivia D S. Data Analysis, a Bayesian Tutorial Oxford: Clarendon Press, 1996

[86] Berger J. O. Statistical Decision Theory. Springer-Verlag, 1980

[87] Raiffa H , R. Schlaifer. Applied Statistical Decision Theory. Boston: Harvard University, 1961

[88] Box G E P, Tiao G C. Bayesian inference in Statistical Analysis. Roading: Addison-Wesly, 1973

[89] Gelman A, et al. Bayesian data analysis. New York: Chapman-Hall, 1995

[90] Press S J. Bayesian statistics. New York: John Wiley-Sons, 1989

[91] Robbins H. Some thoughts on empirical Bayes estimation. Ann. of Statistics, 1983, 11: 713–723

[92] C. R. 劳. 线性统计推断及其应用. 北京：科学出版社, 1987

[93] 赵选民等. 数理统计. 第二版. 北京：科学出版社, 2002

[94] 张尧庭等. 贝叶斯统计推断. 北京：科学出版社, 1991

[95] 范金城. 梅长林. 数据分析. 北京：科学出版社, 2002

[96] Narsky I. arXiv:hep-ex/0005019 (2000)

[97] Zhu Y S (朱永生). Nucl. Instr. Meth. in Phys. Resear., 2007, (A578): 322

[98] Zhu Y S (朱永生). Chinese Phys., 2008, (C32): 363

[99] Zhu Y S (朱永生): http://www.ihep.ac.cn/lunwen/zhuys/BPOCI/BPOCI.html

数理统计表

[100] Owen D. Handbook of Statistical Tables, Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962

[101] Miller L. Table of percentage points of kolmogor ov statistics, J. Amer Statist. Ass, 1956, 51: 111

[102] Resnikoff G, Lieberman G. Table of the Noncentral T-Distribution. Stanford University Press, 1957

[103] Verdoren L. Extended tables of critical values for Wilcoxon's test statistic. Biometrika, 1963, 50: 177

[104] Pearson E, Hartley H. Biometrika tables for statisticians. Cambridge, 1970, 1972, 1, 2

[105] 中国科学院数学研究所概率统计室. 常用数理统计表. 北京：科学出版社, 1974

蒙特卡罗法

[106] Hammersley J, et al. Monte Carlo Methods. London: Methuen & Co. Ltd., 1967

[107] Buslenko N, et al. The Monte Carlo methods. Oxford: Pergamon Press, 1966

[108] Knuth D. The art of computer programming. Addison-Wesley Publishing Company, Reading, Massachusetts, 1969

[109] James F. Monte Carlo phase space. CERN 68-15, Geneva, 1968

[110] Lautrup B. An adaptive multidimensional integration technique. 2nd Colloquium on Advanced Computing Methods in Theoretical Physics, Marseille, 1971

[111] Dufner A. An adaptive multidimensional numerical integration subroutine. Proc. of the Colloquium on Computational Methods in Theoretical Physics, Marseille, 1970

[112] Zerby C. A Monte Carlo calculation of the response of Gamma ray scintillation counter. Methods in Comput. Phys., 1963, 1: 89

[113] Berger M. Monte Carlo calculation of the penetration and diffusion of fast charged particles. Methods in Comput. Phys., 1963, 1: 135

[114] CERN Program Library, D114 RIWIAD, Geneva, 1996.

[115] 裴鹿成, 张孝泽. 蒙特卡罗方法及其在粒子输运问题中的应用. 北京：科学出版社, 1980

最优化方法

[116] James F. Function minimization. CERN 72-21, Geneva, 1972

[117] Kowalik J, et al. Methods for Unconstrained Optimization Problems. New York: American Elsevier Publishing Co., Inc., 1968

[118] Dixon L. Non-linear Optimization. London: English Universities Press, 1972

[119] Sheppey G. Minimization and curve fltting. CERN 68-5, Geneva, 1968

[120] Davidon W. Variance algorithm for minimization. Comp. J., 1968, 10: 406

[121] Nelder J et al. A simplex method for function minimization. Comp. J., 1965, 7: 308

[122] Rosenbrock H. An automatic method for finding the greatest or least value of a function. Comp. J., 1960, 3: 175

[123] Hook R, Jeeves T. Direct search solution of numerical and statistical problems. J. Assoc. Comp.Mach, 1961, 8: 212

[124] Fletcher R, Reeves C. Function minimization by conjugate gradients. Comp. J., 1964, 7: 149

[125] Powell M. An efficient method for finding the minimum of a function of several vari ables without calculating derivatives. Comp. J., 1964, 7: 155

[126] Broyden C. Quasi-Newton methods and the application to function minimization. Math. Compu., 1967, 21: 368

[127] Evans D. et al. Exact treatment of search statistics. Nucl. Instr. Meth., 1967, 56: 289

[128] Gelfand I, Tsetlin M. The principle of nonlocal search in automatic optimization systems. Soviet Phys. Dokl., 1961, 6: 192

[129] Goldstein A, Price J. On descent from local minima. Math. Comput., 1971, 25: 569

[130] CERN Program Library. Geneva, 1989

[131] 希梅尔布劳 D. 张义荣等译. 实用非线性规划. 北京：科学出版社, 1981

[132] 席少霖, 赵风治. 最优化计算方法. 上海：上海科学技术出版社, 1983

[133] 蔡宣三. 最优化与最优控制. 北京：清华大学出版社, 1982

[134] 南京大学数学系计算数学专业. 最优化方法. 北京：科学出版社, 1984

[135] 邓乃扬等. 无约束最优化计算方法. 北京：科学出版社, 1982

[136] 袁亚湘、孙文瑜编著. 最优化理论与方法. 北京：科学出版社, 1997

去弥散方法

[137] Anykeyev V B, et al. Nucl. Instr. Meth. 1991, A303: 350

[138] Anykeyev V B, et al. Nucl. Instr. Meth. 1992, A322: 280

[139] Belogorlov E A, et al. Nucl. Instr. Meth. 1985, A235: 146

[140] Frieden B R. J. Opt. Soc Am. 1972, 62: 511

[141] Frieden B R. Comput. Graphics Image Process. 1980, 12: 40

[142] Frieden B R. Probability, Statistical Optics and Testing. New York: Springer, 1983

[143] Höcker A, et al. Nucl. Instr. Meth. 1996, A372: 469

[144] Jaynes E T. Prior probabilities, IEEE Trans. Syst. Sci. Cyber. 1968(SSC-4):227

[145] Jaynes E T. Monkeys, Kangaroos and N in Maximum Entropy and Bayesian Methods in Applied Statistics. Justice J H (ed.) Cambridge: Cambridge Univ. Press, 1986, 26

[146] Kullback. Information Theory and Statistics. New York: John Wiley, 1964

[147] Louis A K. Inverse Probl. 1992, 8: 709

[148] Narayan R. Ann. Rev. Astron. Astrophys. 1986, 24: 127

[149] Phillips D L. J. ACM. 1962, 9: 84

[150] Press W H, et al. Numerical Recipes. 2nd ed,. Cambridge: Cambridge Univ. Press, 1992

[151] Roe B P. Probability and Statistics in Experimental Phisics. New York: Springer, 1992

[152] Shannon C E. Bell Sys. Tech. J. 1948(27): 379//Shannon C E. Weaver W. The Mathematical Theory of Communication. Urbana: University of Illinois Press, 1949

[153] Michael Schmelling, Nucl. Instr. Meth. 1994, A340: 400

[154] Skilling J, et al. Algorithms and applications, in Maximum Entropy and Bayesian Methods in Inverse Problems, Smith C R, Grady W T. (ed.), Dordrecht : D. Reidel, 1985, 83

[155] Skilling J. Theory of maximum entropy image reconstruction, in Maximum Entropy and Bayesian Methods in Applied Statistics, Justice J H.(ed), Cambridge: Cambridge Univ. Press, 1986, 156.

[156] Tikhonov A N. Sov Math. 1963, 5: 1035

[157] Tikhonov A N, et al. Solutions of Ill-Posed Problems. New York: John Wiley, 1977

[158] Wahba G. Smoothing and ill-posed problems, in Solution Methods for Integral Equations. New York: Plenum Press, 1979

[159] Zech G. Comparing statistical data to Monte Carlo simulation-parameter fitting and unfolding, DESY 95-113, 1995

[160] Zhigunov V P. Nucl. Instr. Meth. 1983, 216: 183
[161] Zhigunov V P, et al. Nucl. Instr. Meth. 1988, A273: 362

概率和统计在粒子和核物理实验中的应用

[162] Mark-J Collaboration①. Test of quantum electrodynamics at $\sqrt{s} = 13$ and 17 GeV. Phys. Rev. Lett., 1979, 42: 1110
[163] Mark-J Collaboration, Discovery of three jet events and a test of quantum chromodynamics at PETRA. Phys. Rev. Lett., 1979, 43: 830
[164] Mark-J Collaboration. Study of electron-positron collisions at center-of-mass energies of 27.4 and 27.7 GeV at PETRA. Phys. Rev. Lett., 1979, 43: 901
[165] Mark-J Collaboration. Test of universality of charged leptons. Phys. Rev. Lett., 1979, 43: 1915
[166] Mark-J Collaboration. Search for the production of a new quark flavor at the C. M. system energies between 33 and 35.8 Gev. Phys. Rev. Lett., 1980, 44: 1722
[167] Mark-J Collaboration. Physics with high energy electron positron colliding beams with the Mark-J detector. Phys. Rep., 1980, 63: 337
[168] Mark-J Collaboration. Search for top quark and a test of models without top quark at the highest PETRA energies. Phys. Rev. Lett., 1983, 50: 799
[169] Mark-J Collaboration. A summary of experimental results from Mark-J: high energy e^+e^t collisions at PETRA. Phys. Rep., 1984, 109: 131
[170] Zhu Y S (朱永生). Monte Carlo calculation on achievable accuracy of Tau lifetime measurement by using TEC type vertex chamber in e^+e^t interaction at center of mass energy $\sqrt{s} = 40$ GeV. Chinese Physics, 1985, 5: 100
[171] Mark-J 合作组. 光子、轻子、层子及胶子. 中国科学, 1980
[172] Mark-J 合作组. 高能正负电子对撞物理的最新实验结果 ——Mark-J 实验总结. 物理学进展, 1984,(4): 301
[173] 朱永生. 质心系能量 40GeV e^+e^- 反应中, 利用时间扩展室测定 τ 粒子寿命可达到精度的蒙特卡罗计算. 高能物理与核物理, 1984, 5: 513
[174] BES Collaboration②. Measurement of the mass of the Tau lepton. Phys. Rev. Lett., 1992, 69: 3021
[175] BES Collaboration. A Measurement of J/ψ decay widths. Phys. Lett., 1995, B355: 374
[176] BES Collaboration. Measurement of the cross section for $e^+e^- \to$Hadrons at center-of-mass energies from 2～5 GeV. Phys. Rev. Lett., 2002, 88: 101802

① Mark-J Collaboration(Mark-J 合作组) 是丁肇中教授领导的粒子物理实验研究组. 本书作者是该组成员之一.

② BES Collaboration (BES 合作组) 是中国高能物理研究所和美国、日本等国内外近 20 个研究所和大学组成的国际粒子物理实验研究组. 本书作者是该组成员之一.

[177] BES Collaboration. A measurement of $\psi(2S)$ resonance parameters. Phys. Lett., 2002, B550: 24

[178] BES Collaboration. $\psi(2S)$ two-and three-body hadronic decays. Phys. Rev., 2003, D67: 052002

[179] BES Collaboration. Measurement of $\psi(2S)$ decays into vector-tensor final states. Phys. Rev., 2004, D69: 072001

[180] BES Collaboration. $\psi(2S)$ decays into J/ψ plus two photons. Phys. Rev., 2004, D70: 012006

[181] BES Collaboration. Measurements of $\psi(2S)$ decays into $\varphi\pi^0$, $\varphi\eta$, $\varphi\eta\prime$, $\omega\eta$, and $\omega\eta\prime$. Phys. Rev., 2004, D70: 112003

[182] BES Collaboration. Measurement of the final states $\omega\pi^0$, $\rho\eta$, and $\rho\eta\prime$ from $\psi(2S)$ electromagnetic decays and e^+e^- annihilations. Phys. Rev., 2004, D70: 112007

[183] Zhu Y S (朱永生). Statistics in HEP Data Analysis, in “Physics at BES-III”. (ed. by Kuang-Ta Chao and Yifang Wang) Int. J. Mod. Phys. , 2009, A24 Suppl. 1. Appendix A

[184] Olivero J J, et al. J. Quant. Spectrosc. Radiat. Transfer 1977, 17: 233

[185] Bruce S D, et al. J. Magn. Reson. 2000, 142: 57

[186] Mclean A B, et al. J. Electron Spectrosc. Relat. Phenom. 1994, 69: 125; Puerta J, et al. Appl. Opt. 1981, 20: 3923

[187] Wertheim G K, et al. Rev. Sci. Instrum. 1974, 45: 1369

附　　表

表 1　二项分布表

$$B(r;n,p)=\binom{n}{r}p^r(1-p)^{n-r}$$

表中只列出 $p \leqslant 0.50$ 的二项分布概率值; $p > 0.50$ 的概率值可由下面的关系式求出：

$$B(r;n,p)=B(n-r;n,1-p).$$

n	r \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
1	0	0.9900	0.9800	0.9700	0.9500	0.9000	0.8500	0.8000	0.7500	0.7000	0.6000	0.5000
	1	0.0100	0.0200	0.0300	0.0500	0.1000	0.1500	0.2000	0.2500	0.3000	0.4000	0.5000
2	0	0.9801	0.9604	0.9409	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.3600	0.2500
	1	0.0198	0.0392	0.0582	0.0950	0.1800	0.2550	0.3200	0.3750	0.4200	0.4800	0.5000
	2	0.0001	0.0004	0.0009	0.0025	0.0100	0.0225	0.0400	0.0625	0.0900	0.1600	0.2500
3	0	0.9703	0.9412	0.9127	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2160	0.1250
	1	0.0294	0.0576	0.0847	0.1354	0.2430	0.3251	0.3840	0.4219	0.4410	0.4320	0.3750
	2	0.0003	0.0012	0.0026	0.0071	0.0270	0.0574	0.0960	0.1406	0.1890	0.2880	0.3750
	3	0.0000	0.0000	0.0000	0.0001	0.0010	0.0034	0.0080	0.0156	0.0270	0.0640	0.1250
4	0	0.9606	0.9224	0.8853	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1296	0.0625
	1	0.0388	0.0753	0.1095	0.1715	0.2916	0.3685	0.4096	0.4219	0.4116	0.3456	0.2500
	2	0.0006	0.0023	0.0051	0.0135	0.0486	0.0975	0.1536	0.2109	0.2646	0.3456	0.3750
	3	0.0000	0.0000	0.0001	0.0005	0.0036	0.0115	0.0256	0.0469	0.0756	0.1536	0.2500
	4	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0016	0.0039	0.0081	0.0256	0.0625
5	0	0.9510	0.9039	0.8587	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.0778	0.0313
	1	0.0480	0.0922	0.1328	0.2036	0.3281	0.3915	0.4096	0.3955	0.3601	0.2592	0.1563
	2	0.0010	0.0038	0.0082	0.0214	0.0729	0.1382	0.2048	0.2637	0.3087	0.3456	0.3125
	3	0.0000	0.0001	0.0003	0.0011	0.0081	0.0244	0.0512	0.0879	0.1323	0.2304	0.3125
	4	0.0000	0.0000	0.0000	0.0000	0.0005	0.0022	0.0064	0.0146	0.0284	0.0768	0.1563
	5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0.0024	0.0102	0.0313
6	0	0.9415	0.8858	0.8330	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0467	0.0156

续表

n	r \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
	1	0.0571	0.1085	0.1546	0.2321	0.3543	0.3993	0.3932	0.3560	0.3025	0.1866	0.0937
	2	0.0014	0.0055	0.0120	0.0305	0.0984	0.1762	0.2458	0.2966	0.3241	0.3110	0.2344
	3	0.0000	0.0002	0.0005	0.0021	0.0146	0.0415	0.0819	0.1318	0.1852	0.2765	0.3125
	4	0.0000	0.0000	0.0000	0.0001	0.0012	0.0055	0.0154	0.0330	0.0595	0.1382	0.2344
	5	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0015	0.0044	0.0102	0.0369	0.0937
	6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0007	0.0041	0.0156
7	0	0.9321	0.8681	0.8080	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0280	0.0078
	1	0.0659	0.1240	0.1749	0.2573	0.3720	0.3960	0.3670	0.3115	0.2471	0.1306	0.0547
	2	0.0020	0.0076	0.0162	0.0406	0.1240	0.2097	0.2753	0.3115	0.3177	0.2613	0.1641
	3	0.0000	0.0003	0.0008	0.0036	0.0230	0.0617	0.1147	0.1730	0.2269	0.2903	0.2734
	4	0.0000	0.0000	0.0000	0.0002	0.0026	0.0109	0.0287	0.0577	0.0972	0.1935	0.2734
	5	0.0000	0.0000	0.0000	0.0000	0.0002	0.0012	0.0043	0.0115	0.0250	0.0774	0.1641
	6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0013	0.0036	0.0172	0.0547
	7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0016	0.0078
8	0	0.9227	0.8508	0.7837	0.6634	0.4305	0.2725	0.1678	0.1001	0.0576	0.0168	0.0039
	1	0.0746	0.1389	0.1939	0.2793	0.3826	0.3847	0.3355	0.2670	0.1977	0.0896	0.0313
	2	0.0026	0.0099	0.0210	0.0515	0.1488	0.2376	0.2936	0.3115	0.2965	0.2090	0.1094
	3	0.0001	0.0004	0.0013	0.0054	0.0331	0.0839	0.1468	0.2076	0.2541	0.2787	0.2188
	4	0.0000	0.0000	0.0001	0.0004	0.0046	0.0185	0.0459	0.0865	0.1361	0.2322	0.2734
	5	0.0000	0.0000	0.0000	0.0000	0.0004	0.0026	0.0092	0.0231	0.0467	0.1239	0.2188
	6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011	0.0038	0.0100	0.0413	0.1094
	7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0012	0.0079	0.0313
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0007	0.0039
9	0	0.9135	0.8337	0.7602	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0101	0.0020
	1	0.0830	0.1531	0.2116	0.2985	0.3874	0.3679	0.3020	0.2253	0.1556	0.0605	0.0176
	2	0.0034	0.0125	0.0262	0.0629	0.1722	0.2597	0.3020	0.3003	0.2668	0.1612	0.0703
	3	0.0001	0.0006	0.0019	0.0077	0.0446	0.1069	0.1762	0.2336	0.2668	0.2508	0.1641
	4	0.0000	0.0000	0.0001	0.0006	0.0074	0.0283	0.0661	0.1168	0.1715	0.2508	0.2461
	5	0.0000	0.0000	0.0000	0.0000	0.0008	0.0050	0.0165	0.0389	0.0735	0.1672	0.2461
	6	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0028	0.0087	0.0210	0.0743	0.1641
	7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0012	0.0039	0.0212	0.0703
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0035	0.0176
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0020
10	0	0.9044	0.8171	0.7374	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0060	0.0010
	1	0.0914	0.1667	0.2281	0.3151	0.3874	0.3474	0.2684	0.1877	0.1211	0.0403	0.0098
	2	0.0042	0.0153	0.0317	0.0746	0.1937	0.2759	0.3020	0.2816	0.2335	0.1209	0.0439

续表

	p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
n	r											
	3	0.0001	0.0008	0.0026	0.0105	0.0574	0.1298	0.2013	0.2503	0.2668	0.2150	0.1172
	4	0.0000	0.0000	0.0001	0.0010	0.0112	0.0401	0.0881	0.1460	0.2001	0.2508	0.2051
	5	0.0000	0.0000	0.0000	0.0001	0.0015	0.0085	0.0264	0.0584	0.1029	0.2007	0.2461
	6	0.0000	0.0000	0.0000	0.0000	0.0001	0.0012	0.0055	0.0162	0.0368	0.1115	0.2051
	7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0008	0.0031	0.0090	0.0425	0.1172
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0014	0.0106	0.0439
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0016	0.0098
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0010
11	0	0.8953	0.8007	0.7153	0.5688	0.3138	0.1673	0.0859	0.0422	0.0198	0.0036	0.0005
	1	0.0995	0.1798	0.2433	0.3293	0.3835	0.3248	0.2362	0.1549	0.0932	0.0266	0.0054
	2	0.0050	0.0183	0.0376	0.0867	0.2131	0.2866	0.2953	0.2581	0.1998	0.0887	0.0269
	3	0.0002	0.0011	0.0035	0.0137	0.0710	0.1517	0.2215	0.2581	0.2568	0.1774	0.0806
	4	0.0000	0.0000	0.0002	0.0014	0.0158	0.0536	0.1107	0.1721	0.2201	0.2365	0.1611
	5	0.0000	0.0000	0.0000	0.0001	0.0025	0.0132	0.0388	0.0803	0.1321	0.2207	0.2256
	6	0.0000	0.0000	0.0000	0.0000	0.0003	0.0023	0.0097	0.0268	0.0566	0.1471	0.2256
	7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0017	0.0064	0.0173	0.0701	0.1611
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011	0.0037	0.0234	0.0806
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0052	0.0269
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0007	0.0054
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005
12	0	0.8864	0.7847	0.6938	0.5404	0.2824	0.1422	0.0687	0.0317	0.0138	0.0022	0.0002
	1	0.1074	0.1922	0.2575	0.3413	0.3766	0.3012	0.2062	0.1267	0.0712	0.0174	0.0029
	2	0.0060	0.0216	0.0438	0.0988	0.2301	0.2924	0.2835	0.2323	0.1678	0.0639	0.0161
	3	0.0002	0.0015	0.0045	0.0173	0.0852	0.1720	0.2362	0.2581	0.2397	0.1419	0.0537
	4	0.0000	0.0001	0.0003	0.0021	0.0213	0.0683	0.1329	0.1936	0.2311	0.2128	0.1208
	5	0.0000	0.0000	0.0000	0.0002	0.0038	0.0193	0.0532	0.1032	0.1585	0.2270	0.1934
	6	0.0000	0.0000	0.0000	0.0000	0.0005	0.0040	0.0155	0.0401	0.0792	0.1766	0.2256
	7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0033	0.0115	0.0291	0.1009	0.1934
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0024	0.0078	0.0420	0.1208
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004	0.0015	0.0125	0.0537
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0025	0.0161
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0029
	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002
13	0	0.8775	0.7690	0.6730	0.5133	0.2542	0.1209	0.0550	0.0238	0.0097	0.0013	0.0001
	1	0.1152	0.2040	0.2706	0.3512	0.3672	0.2774	0.1787	0.1029	0.0540	0.0113	0.0016
	2	0.0070	0.0250	0.0502	0.1109	0.2448	0.2937	0.2680	0.2059	0.1388	0.0453	0.0095
	3	0.0003	0.0019	0.0057	0.0214	0.0997	0.1900	0.2457	0.2517	0.2181	0.1107	0.0349

续表

n	r \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
	4	0.0000	0.0001	0.0004	0.0028	0.0277	0.0838	0.1535	0.2097	0.2337	0.1845	0.0873
	5	0.0000	0.0000	0.0000	0.0003	0.0055	0.0266	0.0691	0.1258	0.1803	0.2214	0.1571
	6	0.0000	0.0000	0.0000	0.0000	0.0008	0.0063	0.0230	0.0559	0.1030	0.1968	0.2095
	7	0.0000	0.0000	0.0000	0.0000	0.0001	0.0011	0.0058	0.0186	0.0442	0.1312	0.2095
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0011	0.0047	0.0142	0.0656	0.1571
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0009	0.0034	0.0243	0.0873
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0065	0.0349
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0012	0.0095
	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0016
	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
14	0	0.8687	0.7536	0.6528	0.4877	0.2288	0.1028	0.0440	0.0178	0.0068	0.0008	0.0001
	1	0.1229	0.2153	0.2827	0.3593	0.3559	0.2539	0.1539	0.0832	0.0407	0.0073	0.0009
	2	0.0081	0.0286	0.0568	0.1229	0.2570	0.2912	0.2501	0.1802	0.1134	0.0317	0.0056
	3	0.0003	0.0023	0.0070	0.0259	0.1142	0.2056	0.2501	0.2402	0.1943	0.0845	0.0222
	4	0.0000	0.0001	0.0006	0.0037	0.0349	0.0998	0.1720	0.2202	0.2290	0.1549	0.0611
	5	0.0000	0.0000	0.0000	0.0004	0.0078	0.0352	0.0860	0.1468	0.1963	0.2066	0.1222
	6	0.0000	0.0000	0.0000	0.0000	0.0013	0.0093	0.0322	0.0734	0.1262	0.2066	0.1833
	7	0.0000	0.0000	0.0000	0.0000	0.0002	0.0019	0.0092	0.0280	0.0618	0.1574	0.2095
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0020	0.0082	0.0232	0.0918	0.1833
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0018	0.0066	0.0408	0.1222
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0014	0.0136	0.0611
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0033	0.0222
	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0056
	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0009
	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
15	0	0.8601	0.7386	0.6333	0.4633	0.2059	0.0874	0.0352	0.0134	0.0047	0.0005	0.0000
	1	0.1303	0.2261	0.2938	0.3658	0.3432	0.2312	0.1319	0.0668	0.0305	0.0047	0.0005
	2	0.0092	0.0323	0.0636	0.1348	0.2669	0.2856	0.2309	0.1559	0.0916	0.0219	0.0032
	3	0.0004	0.0029	0.0085	0.0307	0.1285	0.2184	0.2501	0.2252	0.1700	0.0634	0.0139
	4	0.0000	0.0002	0.0008	0.0049	0.0428	0.1156	0.1876	0.2252	0.2186	0.1268	0.0417
	5	0.0000	0.0000	0.0001	0.0006	0.0105	0.0449	0.1032	0.1651	0.2061	0.1859	0.0916
	6	0.0000	0.0000	0.0000	0.0000	0.0019	0.0132	0.0430	0.0917	0.1472	0.2066	0.1527
	7	0.0000	0.0000	0.0000	0.0000	0.0003	0.0030	0.0138	0.0393	0.0811	0.1771	0.1964
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0035	0.0131	0.0348	0.1181	0.1964
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0007	0.0034	0.0116	0.0612	0.1527
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0007	0.0030	0.0245	0.0916
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0074	0.0417
	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0016	0.0139

续表

n	r \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0032
	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005
	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
16	0	0.8515	0.7238	0.6143	0.4401	0.1853	0.0743	0.0281	0.0100	0.0033	0.0003	0.0000
	1	0.1376	0.2363	0.3040	0.3706	0.3294	0.2097	0.1126	0.0535	0.0228	0.0030	0.0002
	2	0.0104	0.0362	0.0705	0.1463	0.2745	0.2775	0.2111	0.1336	0.0732	0.0150	0.0018
	3	0.0005	0.0034	0.0102	0.0359	0.1423	0.2285	0.2463	0.2079	0.1465	0.0468	0.0085
	4	0.0000	0.0002	0.0010	0.0061	0.0514	0.1311	0.2001	0.2252	0.2040	0.1014	0.0278
	5	0.0000	0.0000	0.0001	0.0008	0.0137	0.0555	0.1201	0.1802	0.2099	0.1623	0.0667
	6	0.0000	0.0000	0.0000	0.0001	0.0028	0.0180	0.0550	0.1101	0.1649	0.1983	0.1222
	7	0.0000	0.0000	0.0000	0.0000	0.0004	0.0045	0.0197	0.0524	0.1010	0.1889	0.1746
	8	0.0000	0.0000	0.0000	0.0000	0.0001	0.0009	0.0055	0.0197	0.0487	0.1417	0.1964
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0012	0.0058	0.0185	0.0840	0.1746
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0014	0.0056	0.0392	0.1222
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0013	0.0142	0.0667
	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0040	0.0278
	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0008	0.0085
	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0018
	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002
	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
17	0	0.8429	0.7093	0.5958	0.4181	0.1668	0.0631	0.0225	0.0075	0.0023	0.0002	0.0000
	1	0.1447	0.2461	0.3133	0.3741	0.3150	0.1893	0.0957	0.0426	0.0169	0.0019	0.0001
	2	0.0117	0.0402	0.0775	0.1575	0.2800	0.2673	0.1914	0.1136	0.0581	0.0102	0.0010
	3	0.0006	0.0041	0.0120	0.0415	0.1556	0.2359	0.2393	0.1893	0.1245	0.0341	0.0052
	4	0.0000	0.0003	0.0013	0.0076	0.0605	0.1457	0.2093	0.2209	0.1868	0.0796	0.0182
	5	0.0000	0.0000	0.0001	0.0010	0.0175	0.0668	0.1361	0.1914	0.2081	0.1379	0.0472
	6	0.0000	0.0000	0.0000	0.0001	0.0039	0.0236	0.0680	0.1276	0.1784	0.1839	0.0944
	7	0.0000	0.0000	0.0000	0.0000	0.0007	0.0065	0.0267	0.0668	0.1201	0.1927	0.1484
	8	0.0000	0.0000	0.0000	0.0000	0.0001	0.0014	0.0084	0.0279	0.0644	0.1606	0.1855
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0021	0.0093	0.0276	0.1070	0.1855
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0025	0.0095	0.0571	0.1484
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0026	0.0242	0.0944
	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0006	0.0081	0.0472
	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0021	0.0182
	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0052
	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0010
	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

续表

n	r \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
18	0	0.8345	0.6951	0.5780	0.3972	0.1501	0.0536	0.0180	0.0056	0.0016	0.0001	0.0000
	1	0.1517	0.2554	0.3217	0.3763	0.3002	0.1704	0.0811	0.0338	0.0126	0.0012	0.0001
	2	0.0130	0.0443	0.0846	0.1683	0.2835	0.2556	0.1723	0.0958	0.0458	0.0069	0.0006
	3	0.0007	0.0048	0.0140	0.0473	0.1680	0.2406	0.2297	0.1704	0.1046	0.0246	0.0031
	4	0.0000	0.0004	0.0016	0.0093	0.0700	0.1592	0.2153	0.2130	0.1681	0.0614	0.0117
	5	0.0000	0.0000	0.0001	0.0014	0.0218	0.0787	0.1507	0.1988	0.2017	0.1146	0.0327
	6	0.0000	0.0000	0.0000	0.0002	0.0052	0.0301	0.0816	0.1436	0.1873	0.1655	0.0708
	7	0.0000	0.0000	0.0000	0.0000	0.0010	0.0091	0.0350	0.0820	0.1376	0.1892	0.1214
	8	0.0000	0.0000	0.0000	0.0000	0.0002	0.0022	0.0120	0.0376	0.0811	0.1734	0.1669
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0033	0.0139	0.0386	0.1284	0.1855
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0008	0.0042	0.0149	0.0771	0.1669
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0010	0.0046	0.0374	0.1214
	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0012	0.0145	0.0708
	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0045	0.0327
	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011	0.0117
	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0031
	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006
	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
	18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
19	0	0.8262	0.6812	0.5606	0.3774	0.1351	0.0456	0.0144	0.0042	0.0011	0.0001	0.0000
	1	0.1586	0.2642	0.3294	0.3774	0.2852	0.1529	0.0685	0.0268	0.0093	0.0008	0.0000
	2	0.0144	0.0485	0.0917	0.1787	0.2852	0.2428	0.1540	0.0803	0.0358	0.0046	0.0003
	3	0.0008	0.0056	0.0161	0.0533	0.1796	0.2428	0.2182	0.1517	0.0869	0.0175	0.0018
	4	0.0000	0.0005	0.0020	0.0112	0.0798	0.1714	0.2182	0.2023	0.1491	0.0467	0.0074
	5	0.0000	0.0000	0.0002	0.0018	0.0266	0.0907	0.1636	0.2023	0.1916	0.0933	0.0222
	6	0.0000	0.0000	0.0000	0.0002	0.0069	0.0374	0.0955	0.1574	0.1916	0.1451	0.0518
	7	0.0000	0.0000	0.0000	0.0000	0.0014	0.0122	0.0443	0.0974	0.1525	0.1797	0.0961
	8	0.0000	0.0000	0.0000	0.0000	0.0002	0.0032	0.0166	0.0487	0.0981	0.1797	0.1442
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0007	0.0051	0.0198	0.0514	0.1464	0.1762
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0013	0.0066	0.0220	0.0976	0.1762
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0018	0.0077	0.0532	0.1442
	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0022	0.0237	0.0961
	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0005	0.0085	0.0518
	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0024	0.0222
	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0074
	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0018
	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003
	18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

续表

n	r \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
20	0	0.8179	0.6676	0.5438	0.2585	0.1216	0.0388	0.0115	0.0032	0.0008	0.0000	0.0000
	1	0.1652	0.2725	0.3364	0.3774	0.2702	0.1368	0.0576	0.0211	0.0068	0.0005	0.0000
	2	0.0159	0.0528	0.0988	0.1887	0.2852	0.2293	0.1369	0.0669	0.0278	0.0031	0.0002
	3	0.0010	0.0065	0.0183	0.0596	0.1901	0.2428	0.2054	0.1339	0.0716	0.0123	0.0011
	4	0.0000	0.0006	0.0024	0.0133	0.0898	0.1821	0.2182	0.1897	0.1304	0.0350	0.0046
	5	0.0000	0.0000	0.0002	0.0022	0.0319	0.1028	0.1746	0.2023	0.1789	0.0746	0.0148
	6	0.0000	0.0000	0.0000	0.0003	0.0089	0.0454	0.1091	0.1686	0.1916	0.1244	0.0370
	7	0.0000	0.0000	0.0000	0.0000	0.0020	0.0160	0.0545	0.1124	0.1643	0.1659	0.0739
	8	0.0000	0.0000	0.0000	0.0000	0.0004	0.0046	0.0222	0.0609	0.1144	0.1797	0.1201
	9	0.0000	0.0000	0.0000	0.0000	0.0001	0.0011	0.0074	0.0271	0.0654	0.1597	0.1602
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0020	0.0099	0.0308	0.1171	0.1762
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0030	0.0120	0.0710	0.1602
	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0008	0.0039	0.0355	0.1201
	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0010	0.0146	0.0739
	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0049	0.0370
	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0013	0.0148
	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0046
	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011
	18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002
	19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	20	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
25	0	0.7778	0.6035	0.4670	0.2774	0.0718	0.0172	0.0038	0.0008	0.0001	0.0000	0.0000
	1	0.1964	0.3079	0.3611	0.3650	0.1994	0.0759	0.0236	0.0063	0.0014	0.0000	0.0000
	2	0.0238	0.0754	0.1340	0.2305	0.2659	0.1607	0.0708	0.0251	0.0074	0.0004	0.0000
	3	0.0018	0.0118	0.0318	0.0830	0.2265	0.2174	0.1358	0.0641	0.0243	0.0019	0.0001
	4	0.0001	0.0013	0.0054	0.0269	0.1384	0.2110	0.1867	0.1175	0.0572	0.0071	0.0004
	5	0.0000	0.0001	0.0007	0.0060	0.0646	0.1564	0.1960	0.1645	0.1030	0.0199	0.0016
	6	0.0000	0.0000	0.0001	0.0010	0.0239	0.0920	0.1633	0.1828	0.1472	0.0442	0.0053
	7	0.0000	0.0000	0.0000	0.0001	0.0072	0.0441	0.1108	0.1654	0.1712	0.0800	0.0143
	8	0.0000	0.0000	0.0000	0.0000	0.0018	0.0175	0.0623	0.1241	0.1651	0.1200	0.0322
	9	0.0000	0.0000	0.0000	0.0000	0.0004	0.0058	0.0294	0.0781	0.1336	0.1511	0.0609
	10	0.0000	0.0000	0.0000	0.0000	0.0001	0.0016	0.0118	0.0417	0.0916	0.1612	0.0974
	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0040	0.0189	0.0536	0.1465	0.1328
	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0012	0.0074	0.0268	0.1140	0.1550
	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0025	0.0115	0.0760	0.1550
	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0007	0.0042	0.0434	0.1328
	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0013	0.0212	0.0974
	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0088	0.0609
	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0031	0.0322

续表

n	r \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
	18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0009	0.0143
	19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0053
	20	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0016
	21	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004
	22	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
	23	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	24	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	25	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
30	0	0.7397	0.5455	0.4010	0.2146	0.0424	0.0076	0.0012	0.0002	0.0000	0.0000	0.0000
	1	0.2242	0.3340	0.3721	0.3389	0.1413	0.0404	0.0093	0.0018	0.0003	0.0000	0.0000
	2	0.0328	0.0988	0.1669	0.2586	0.2277	0.1034	0.0337	0.0086	0.0018	0.0000	0.0000
	3	0.0031	0.0188	0.0482	0.1270	0.2361	0.1703	0.0785	0.0269	0.0072	0.0003	0.0000
	4	0.0002	0.0026	0.0101	0.0451	0.1771	0.2028	0.1325	0.0604	0.0208	0.0012	0.0000
	5	0.0000	0.0003	0.0016	0.0124	0.1023	0.1861	0.1723	0.1047	0.0464	0.0041	0.0001
	6	0.0000	0.0000	0.0002	0.0027	0.0474	0.1368	0.1795	0.1455	0.0829	0.0115	0.0006
	7	0.0000	0.0000	0.0000	0.0005	0.0180	0.0828	0.1538	0.1662	0.1219	0.0263	0.0019
	8	0.0000	0.0000	0.0000	0.0001	0.0058	0.0420	0.1106	0.1593	0.1501	0.0505	0.0055
	9	0.0000	0.0000	0.0000	0.0000	0.0016	0.0181	0.0676	0.1298	0.1573	0.0823	0.0133
	10	0.0000	0.0000	0.0000	0.0000	0.0004	0.0067	0.0355	0.0909	0.1416	0.1152	0.0280
	11	0.0000	0.0000	0.0000	0.0000	0.0001	0.0022	0.0161	0.0551	0.1103	0.1396	0.0509
	12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0064	0.0291	0.0749	0.1474	0.0806
	13	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0022	0.0134	0.0444	0.1360	0.1115
	14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0007	0.0054	0.0231	0.1101	0.1354
	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0019	0.0106	0.0783	0.1445
	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0042	0.0489	0.1354
	17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0015	0.0269	0.1115
	18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0129	0.0806
	19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0054	0.0509
	20	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0020	0.0280
	21	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0133
	22	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0055
	23	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0019
	24	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006
	25	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
	26	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	27	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	28	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	29	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	30	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

表 2　累积二项分布表

$$F(x;\ \ n,\ \ p)=\sum_{r=0}^{x}B(r;\ \ n,\ \ p)$$

表中只列出 $p\leqslant 0.50$ 的累积二项分布 $F(x;n,p)$ 值; $p>0.50$ 的累积分布可由下面的关系式求出：

$$F(x;n,p)=1-F(n-x-1;n,1-p),\quad 0\leqslant x\leqslant n-1.$$

n	x \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
1	0	0.9900	0.9800	0.9700	0.9500	0.9000	0.8500	0.8000	0.7500	0.7000	0.6000	0.5000
	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0	0.9801	0.9604	0.9409	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.3600	0.2500
	1	0.9999	0.9996	0.9991	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8400	0.7500
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	0.9703	0.9412	0.9127	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2160	0.1250
	1	0.9997	0.9988	0.9974	0.9927	0.9720	0.9392	0.8960	0.8438	0.7840	0.6480	0.5000
	2	1.0000	1.0000	1.0000	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9360	0.8750
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	0.9606	0.9224	0.8853	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1296	0.0625
	1	0.9994	0.9977	0.9948	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.4752	0.3125
	2	1.0000	1.0000	0.9999	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8208	0.6875
	3	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9744	0.9375
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	0.9510	0.9039	0.8587	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.0778	0.0313
	1	0.9990	0.9962	0.9915	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.3370	0.1875
	2	1.0000	0.9999	0.9997	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.6826	0.5000
	3	1.0000	1.0000	1.0000	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9130	0.8125
	4	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990	0.9976	0.9898	0.9687
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	0.9415	0.8858	0.8330	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0467	0.0156
	1	0.9985	0.9943	0.9875	0.9672	0.8857	0.7765	0.6554	0.5339	0.4202	0.2333	0.1094
	2	1.0000	0.9998	0.9995	0.9978	0.9841	0.9527	0.9011	0.8306	0.7443	0.5443	0.3438
	3	1.0000	1.0000	1.0000	0.9999	0.9987	0.9941	0.9830	0.9624	0.9295	0.8208	0.6562
	4	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9954	0.9891	0.9590	0.8906

续表

	p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
n	x											
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9959	0.9844
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	0	0.9321	0.8681	0.8080	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0280	0.0078
	1	0.9980	0.9921	0.9829	0.9556	0.8503	0.7166	0.5767	0.4449	0.3294	0.1586	0.0625
	2	1.0000	0.9997	0.9991	0.9962	0.9743	0.9262	0.8520	0.7564	0.6471	0.4199	0.2266
	3	1.0000	1.0000	1.0000	0.9998	0.9973	0.9879	0.9667	0.9294	0.8740	0.7102	0.5000
	4	1.0000	1.0000	1.0000	1.0000	0.9998	0.9988	0.9953	0.9871	0.9712	0.9037	0.7734
	5	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9987	0.9962	0.9812	0.9375
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9984	0.9922
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8	0	0.9227	0.8508	0.7837	0.6634	0.4305	0.2725	0.1678	0.1001	0.0576	0.0168	0.0039
	1	0.9973	0.9897	0.9777	0.9428	0.8131	0.6572	0.5033	0.3671	0.2553	0.1064	0.0352
	2	0.9999	0.9996	0.9987	0.9942	0.9619	0.8948	0.7969	0.6785	0.5518	0.3154	0.1445
	3	1.0000	1.0000	0.9999	0.9996	0.9950	0.9786	0.9437	0.8862	0.8059	0.5941	0.3633
	4	1.0000	1.0000	1.0000	1.0000	0.9996	0.9971	0.9896	0.9727	0.9420	0.8263	0.6367
	5	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9988	0.9958	0.9887	0.9502	0.8555
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9987	0.9915	0.9648
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9961
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	0	0.9135	0.8337	0.7602	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0101	0.0020
	1	0.9966	0.9869	0.9718	0.9288	0.7748	0.5995	0.4362	0.3003	0.1960	0.0705	0.0195
	2	0.9999	0.9994	0.9980	0.9916	0.9470	0.8591	0.7382	0.6007	0.4628	0.2318	0.0898
	3	1.0000	1.0000	0.9999	0.9994	0.9917	0.9661	0.9144	0.8343	0.7297	0.4826	0.2539
	4	1.0000	1.0000	1.0000	1.0000	0.9991	0.9944	0.9804	0.9511	0.9012	0.7334	0.5000
	5	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9969	0.9900	0.9747	0.9006	0.7461
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9987	0.9957	0.9750	0.9102
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9962	0.9805
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9980
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	0	0.9044	0.8171	0.7374	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0060	0.0010
	1	0.9957	0.9838	0.9655	0.9139	0.7361	0.5443	0.3758	0.2440	0.1493	0.0464	0.0107
	2	0.9999	0.9991	0.9972	0.9885	0.9298	0.8202	0.6778	0.5256	0.3828	0.1673	0.0547
	3	1.0000	1.0000	0.9999	0.9990	0.9872	0.9500	0.8791	0.7759	0.6496	0.3823	0.1719
	4	1.0000	1.0000	1.0000	0.9999	0.9984	0.9901	0.9672	0.9219	0.8497	0.6331	0.3770
	5	1.0000	1.0000	1.0000	1.0000	0.9999	0.9986	0.9936	0.9803	0.9527	0.8338	0.6230
	6	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9991	0.9965	0.9894	0.9452	0.8281

续表

n	x \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9877	0.9453
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9983	0.9893
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9990
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
11	0	0.8953	0.8007	0.7153	0.5688	0.3138	0.1673	0.0859	0.0422	0.0198	0.0036	0.0005
	1	0.9948	0.9805	0.9587	0.8981	0.6974	0.4922	0.3221	0.1971	0.1130	0.0302	0.0059
	2	0.9998	0.9988	0.9963	0.9848	0.9104	0.7788	0.6174	0.4552	0.3127	0.1189	0.0327
	3	1.0000	1.0000	0.9998	0.9984	0.9815	0.9306	0.8389	0.7133	0.5696	0.2963	0.1133
	4	1.0000	1.0000	1.0000	0.9999	0.9972	0.9841	0.9496	0.8854	0.7897	0.5328	0.2744
	5	1.0000	1.0000	1.0000	1.0000	0.9997	0.9973	0.9883	0.9657	0.9218	0.7535	0.5000
	6	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9980	0.9924	0.9784	0.9006	0.7256
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9988	0.9957	0.9707	0.8867
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9941	0.9673
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9993	0.9941
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
12	0	0.8864	0.7847	0.6938	0.5404	0.2824	0.1422	0.0687	0.0317	0.0138	0.0022	0.0002
	1	0.9938	0.9769	0.9514	0.8816	0.6590	0.4435	0.2749	0.1584	0.0850	0.0196	0.0032
	2	0.9998	0.9985	0.9952	0.9804	0.8891	0.7358	0.5583	0.3907	0.2528	0.0834	0.0193
	3	1.0000	0.9999	0.9997	0.9978	0.9744	0.9078	0.7946	0.6488	0.4925	0.2253	0.0730
	4	1.0000	1.0000	1.0000	0.9998	0.9957	0.9761	0.9274	0.8424	0.7237	0.4382	0.1938
	5	1.0000	1.0000	1.0000	1.0000	0.9995	0.9954	0.9806	0.9456	0.8822	0.6652	0.3872
	6	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9961	0.9857	0.9614	0.8418	0.6128
	7	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9972	0.9905	0.9427	0.8062
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9983	0.9847	0.9270
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9972	0.9807
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9968
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
13	0	0.8775	0.7690	0.6730	0.5133	0.2542	0.1209	0.0550	0.0238	0.0097	0.0013	0.0001
	1	0.9928	0.9730	0.9436	0.8646	0.6213	0.3983	0.2336	0.1267	0.0637	0.0126	0.0017
	2	0.9997	0.9980	0.9938	0.9755	0.8661	0.6920	0.5017	0.3326	0.2025	0.0579	0.0112
	3	1.0000	0.9999	0.9995	0.9969	0.9658	0.8820	0.7473	0.5843	0.4206	0.1686	0.0461
	4	1.0000	1.0000	1.0000	0.9997	0.9935	0.9658	0.9009	0.7940	0.6543	0.3530	0.1334
	5	1.0000	1.0000	1.0000	1.0000	0.9991	0.9925	0.9700	0.9198	0.8346	0.5744	0.2905
	6	1.0000	1.0000	1.0000	1.0000	0.9999	0.9987	0.9930	0.9757	0.9376	0.7712	0.5000
	7	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9988	0.9944	0.9818	0.9023	0.7095

续表

n	x \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9990	0.9960	0.9679	0.8666
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9922	0.9539
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9987	0.9888
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9983
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
14	0	0.8687	0.7536	0.6528	0.4877	0.2288	0.1028	0.0440	0.0178	0.0068	0.0008	0.0001
	1	0.9916	0.9690	0.9355	0.8470	0.5846	0.3567	0.1979	0.1010	0.0475	0.0081	0.0009
	2	0.9997	0.9975	0.9923	0.9699	0.8416	0.6479	0.4481	0.2811	0.1608	0.0398	0.0065
	3	1.0000	0.9999	0.9994	0.9958	0.9559	0.8535	0.6982	0.5213	0.3552	0.1243	0.0287
	4	1.0000	1.0000	1.0000	0.9996	0.9908	0.9533	0.8702	0.7415	0.5842	0.2793	0.0898
	5	1.0000	1.0000	1.0000	1.0000	0.9985	0.9885	0.9561	0.8883	0.7805	0.4859	0.2120
	6	1.0000	1.0000	1.0000	1.0000	0.9998	0.9978	0.9884	0.9617	0.9067	0.6925	0.3953
	7	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9976	0.9897	0.9685	0.8499	0.6047
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9996	0.9978	0.9917	0.9417	0.7880
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9983	0.9825	0.9102
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9961	0.9713
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9994	0.9935
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9991
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
15	0	0.8601	0.7386	0.6333	0.4633	0.2059	0.0874	0.0352	0.0134	0.0047	0.0005	0.0000
	1	0.9904	0.9647	0.9270	0.8290	0.5490	0.3186	0.1671	0.0802	0.0353	0.0052	0.0005
	2	0.9996	0.9970	0.9906	0.9638	0.8159	0.6042	0.3980	0.2361	0.1268	0.0271	0.0037
	3	1.0000	0.9998	0.9992	0.9945	0.9444	0.8227	0.6482	0.4613	0.2969	0.0905	0.0176
	4	1.0000	1.0000	0.9999	0.9994	0.9873	0.9383	0.8358	0.6865	0.5155	0.2173	0.0592
	5	1.0000	1.0000	1.0000	0.9999	0.9978	0.9832	0.9389	0.8516	0.7216	0.4032	0.1509
	6	1.0000	1.0000	1.0000	1.0000	0.9997	0.9964	0.9819	0.9434	0.8689	0.6098	0.3036
	7	1.0000	1.0000	1.0000	1.0000	1.0000	0.9994	0.9958	0.9827	0.9500	0.7869	0.5000
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9992	0.9958	0.9848	0.9050	0.6964
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9992	0.9963	0.9662	0.8491
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9907	0.9408
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9981	0.9824
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9963
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

续表

n	x \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
16	0	0.8515	0.7238	0.6143	0.4401	0.1853	0.0743	0.0281	0.0100	0.0033	0.0003	0.0000
	1	0.9891	0.9601	0.9182	0.8108	0.5147	0.2839	0.1407	0.0635	0.0261	0.0033	0.0003
	2	0.9995	0.9963	0.9887	0.9571	0.7892	0.5614	0.3518	0.1971	0.0994	0.0183	0.0021
	3	1.0000	0.9998	0.9989	0.9930	0.9316	0.7899	0.5981	0.4050	0.2459	0.0651	0.0106
	4	1.0000	1.0000	0.9999	0.9991	0.9830	0.9209	0.7982	0.6302	0.4499	0.1666	0.0384
	5	1.0000	1.0000	1.0000	0.9999	0.9967	0.9765	0.9183	0.8103	0.6598	0.3288	0.1051
	6	1.0000	1.0000	1.0000	1.0000	0.9995	0.9944	0.9733	0.9204	0.8247	0.5272	0.2272
	7	1.0000	1.0000	1.0000	1.0000	0.9999	0.9989	0.9930	0.9729	0.9256	0.7161	0.4018
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9985	0.9925	0.9743	0.8577	0.5982
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9984	0.9929	0.9417	0.7728
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9984	0.9809	0.8949
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9951	0.9616
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9991	0.9894
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9979
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
17	0	0.8429	0.7093	0.5958	0.4181	0.1668	0.0631	0.0225	0.0075	0.0023	0.0002	0.0000
	1	0.9877	0.9554	0.9091	0.7922	0.4818	0.2525	0.1182	0.0501	0.0193	0.0021	0.0001
	2	0.9994	0.9956	0.9866	0.9497	0.7618	0.5198	0.3096	0.1637	0.0774	0.0123	0.0012
	3	1.0000	0.9997	0.9986	0.9912	0.9174	0.7556	0.5489	0.3530	0.2019	0.0464	0.0064
	4	1.0000	1.0000	0.9999	0.9988	0.9779	0.9013	0.7582	0.5739	0.3887	0.1260	0.0245
	5	1.0000	1.0000	1.0000	0.9999	0.9953	0.9681	0.8943	0.7653	0.5968	0.2639	0.0717
	6	1.0000	1.0000	1.0000	1.0000	0.9992	0.9917	0.9623	0.8929	0.7752	0.4478	0.1662
	7	1.0000	1.0000	1.0000	1.0000	0.9999	0.9983	0.9891	0.9598	0.8954	0.6405	0.3145
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9974	0.9876	0.9597	0.8011	0.5000
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9969	0.9873	0.9081	0.6855
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9968	0.9652	0.8338
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9894	0.9283
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9975	0.9755
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9936
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9988
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
18	0	0.8345	0.6951	0.5780	0.3972	0.1501	0.0536	0.0180	0.0056	0.0016	0.0001	0.0000
	1	0.9862	0.9505	0.8997	0.7735	0.4503	0.2241	0.0991	0.0395	0.0142	0.0013	0.0001
	2	0.9993	0.9948	0.9843	0.9419	0.7338	0.4797	0.2713	0.1353	0.0600	0.0082	0.0007

续表

n	x \ p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
	3	1.0000	0.9996	0.9982	0.9891	0.9018	0.7202	0.5010	0.3057	0.1646	0.0328	0.0038
	4	1.0000	1.0000	0.9998	0.9985	0.9718	0.8794	0.7164	0.5187	0.3327	0.0942	0.0154
	5	1.0000	1.0000	1.0000	0.9998	0.9936	0.9581	0.8671	0.7175	0.5344	0.2088	0.0481
	6	1.0000	1.0000	1.0000	1.0000	0.9988	0.9882	0.9487	0.8610	0.7217	0.3743	0.1189
	7	1.0000	1.0000	1.0000	1.0000	0.9998	0.9973	0.9837	0.9431	0.8593	0.5634	0.2403
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9957	0.9807	0.9404	0.7368	0.4073
	9	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9991	0.9946	0.9790	0.8653	0.5927
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9988	0.9939	0.9424	0.7597
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9986	0.9797	0.8811
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9942	0.9519
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9987	0.9846
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9962
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9993
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
19	0	0.8262	0.6812	0.5606	0.3774	0.1351	0.0456	0.0144	0.0042	0.0011	0.0001	0.0000
	1	0.9847	0.9454	0.8900	0.7547	0.4203	0.1985	0.0829	0.0310	0.0104	0.0008	0.0000
	2	0.9991	0.9939	0.9817	0.9335	0.7054	0.4413	0.2369	0.1113	0.0462	0.0055	0.0004
	3	1.0000	0.9995	0.9978	0.9868	0.8850	0.6841	0.4551	0.2631	0.1332	0.0230	0.0022
	4	1.0000	1.0000	0.9998	0.9980	0.9648	0.8556	0.6733	0.4654	0.2822	0.0696	0.0096
	5	1.0000	1.0000	1.0000	0.9998	0.9914	0.9463	0.8369	0.6678	0.4739	0.1629	0.0318
	6	1.0000	1.0000	1.0000	1.0000	0.9983	0.9837	0.9324	0.8251	0.6655	0.3081	0.0835
	7	1.0000	1.0000	1.0000	1.0000	0.9997	0.9959	0.9767	0.9225	0.8180	0.4878	0.1796
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9992	0.9933	0.9713	0.9161	0.6675	0.3238
	9	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9984	0.9911	0.9674	0.8139	0.5000
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9977	0.9895	0.9115	0.6762
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9972	0.9648	0.8204
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9884	0.9165
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9969	0.9682
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9994	0.9904
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9978
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9996
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
20	0	0.8179	0.6676	0.5438	0.3585	0.1216	0.0388	0.0115	0.0032	0.0008	0.0000	0.0000
	1	0.9831	0.9401	0.8802	0.7358	0.3917	0.1756	0.0692	0.0243	0.0076	0.0005	0.0000

续表

	p	0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
n	x											
	2	0.9990	0.9929	0.9790	0.9245	0.6769	0.4049	0.2061	0.0913	0.0355	0.0036	0.0002
	3	1.0000	0.9994	0.9973	0.9841	0.8670	0.6477	0.4114	0.2252	0.1071	0.0160	0.0013
	4	1.0000	1.0000	0.9997	0.9974	0.9568	0.8298	0.6296	0.4148	0.2375	0.0510	0.0059
	5	1.0000	1.0000	1.0000	0.9997	0.9887	0.9327	0.8042	0.6172	0.4164	0.1256	0.0207
	6	1.0000	1.0000	1.0000	1.0000	0.9976	0.9781	0.9133	0.7858	0.6080	0.2500	0.0577
	7	1.0000	1.0000	1.0000	1.0000	0.9996	0.9941	0.9679	0.8982	0.7723	0.4159	0.1316
	8	1.0000	1.0000	1.0000	1.0000	0.9999	0.9987	0.9900	0.9591	0.8867	0.5956	0.2517
	9	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9974	0.9861	0.9520	0.7553	0.4119
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9994	0.9961	0.9829	0.8725	0.5881
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9991	0.9949	0.9435	0.7483
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9987	0.9790	0.8684
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9935	0.9423
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9984	0.9793
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9941
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9987
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
25	0	0.7778	0.6035	0.4670	0.2774	0.0718	0.0172	0.0038	0.0008	0.0001	0.0000	0.0000
	1	0.9742	0.9114	0.8280	0.6424	0.2712	0.0931	0.0274	0.0070	0.0016	0.0001	0.0000
	2	0.9980	0.9868	0.9620	0.8729	0.5371	0.2537	0.0982	0.0321	0.0090	0.0004	0.0000
	3	0.9999	0.9986	0.9938	0.9659	0.7636	0.4711	0.2340	0.0962	0.0332	0.0024	0.0001
	4	1.0000	0.9999	0.9992	0.9928	0.9020	0.6821	0.4207	0.2137	0.0905	0.0095	0.0005
	5	1.0000	1.0000	0.9999	0.9988	0.9666	0.8385	0.6167	0.3783	0.1935	0.0294	0.0020
	6	1.0000	1.0000	1.0000	0.9998	0.9905	0.9305	0.7800	0.5611	0.3407	0.0736	0.0073
	7	1.0000	1.0000	1.0000	1.0000	0.9977	0.9745	0.8909	0.7265	0.5118	0.1536	0.0216
	8	1.0000	1.0000	1.0000	1.0000	0.9995	0.9920	0.9532	0.8506	0.6769	0.2735	0.0539
	9	1.0000	1.0000	1.0000	1.0000	0.9999	0.9979	0.9827	0.9287	0.8106	0.4246	0.1148
	10	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9944	0.9703	0.9022	0.5858	0.2122
	11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9985	0.9893	0.9558	0.7323	0.3450
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9996	0.9966	0.9825	0.8462	0.5000
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9991	0.9940	0.9222	0.6550
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9982	0.9656	0.7878
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9868	0.8852
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9957	0.9461
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9988	0.9784
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9927

续表

n	x	p: 0.01	0.02	0.03	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9980
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995
	21	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	22	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	23	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	24	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
30	0	0.7397	0.5455	0.4010	0.2146	0.0424	0.0076	0.0012	0.0002	0.0000	0.0000	0.0000
	1	0.9639	0.8795	0.7731	0.5535	0.1837	0.0480	0.0105	0.0020	0.0003	0.0000	0.0000
	2	0.9967	0.9783	0.9399	0.8122	0.4114	0.1514	0.0442	0.0106	0.0021	0.0000	0.0000
	3	0.9998	0.9971	0.9881	0.9392	0.6474	0.3217	0.1227	0.0374	0.0093	0.0003	0.0000
	4	1.0000	0.9997	0.9982	0.9844	0.8245	0.5245	0.2552	0.0979	0.0302	0.0015	0.0000
	5	1.0000	1.0000	0.9998	0.9967	0.9268	0.7106	0.4275	0.2026	0.0766	0.0057	0.0002
	6	1.0000	1.0000	1.0000	0.9994	0.9742	0.8474	0.6070	0.3481	0.1595	0.0172	0.0007
	7	1.0000	1.0000	1.0000	0.9999	0.9922	0.9302	0.7608	0.5143	0.2814	0.0435	0.0026
	8	1.0000	1.0000	1.0000	1.0000	0.9980	0.9722	0.8713	0.6736	0.4315	0.0940	0.0081
	9	1.0000	1.0000	1.0000	1.0000	0.9995	0.9903	0.9389	0.8034	0.5888	0.1763	0.0214
	10	1.0000	1.0000	1.0000	1.0000	0.9999	0.9971	0.9744	0.8943	0.7304	0.2915	0.0494
	11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9992	0.9905	0.9493	0.8407	0.4311	0.1002
	12	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9969	0.9784	0.9155	0.5785	0.1808
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9991	0.9918	0.9599	0.7145	0.2923
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9973	0.9831	0.8246	0.4278
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9992	0.9936	0.9029	0.5722
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9979	0.9519	0.7077
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9788	0.8192
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9917	0.8998
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9971	0.9506
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9991	0.9786
	21	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9919
	22	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9974
	23	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9993
	24	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
	25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	26	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	27	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	29	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	30	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

表 3　泊松分布表

$$P(r;\ \mu) = \frac{1}{r!}\ \mu^r \mathrm{e}^{-\mu}$$

μ / r	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066	0.3679
1	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659	0.3679
2	0.0045	0.0164	0.0333	0.0536	0.0758	0.0988	0.1217	0.1438	0.1647	0.1839
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494	0.0613
4	0.0000	0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111	0.0153
5	0.0000	0.0000	0.0000	0.0001	0.0002	0.0004	0.0007	0.0012	0.0020	0.0031
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003	0.0005
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001

μ / r	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	0.3329	0.3012	0.2725	0.2466	0.2231	0.2019	0.1827	0.1653	0.1496	0.1353
1	0.3662	0.3614	0.3543	0.3452	0.3347	0.3230	0.3106	0.2975	0.2842	0.2707
2	0.2014	0.2169	0.2303	0.2417	0.2510	0.2584	0.2640	0.2678	0.2700	0.2707
3	0.0738	0.0867	0.0998	0.1128	0.1255	0.1378	0.1496	0.1607	0.1710	0.1804
4	0.0203	0.0260	0.0324	0.0395	0.0471	0.0551	0.0636	0.0723	0.0812	0.0902
5	0.0045	0.0062	0.0084	0.0111	0.0141	0.0176	0.0216	0.0260	0.0309	0.0361
6	0.0008	0.0012	0.0018	0.0026	0.0035	0.0047	0.0061	0.0078	0.0098	0.0120
7	0.0001	0.0002	0.0003	0.0005	0.0008	0.0011	0.0015	0.0020	0.0027	0.0034
8	0.0000	0.0000	0.0001	0.0001	0.0001	0.0002	0.0003	0.0005	0.0006	0.0009
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0002

μ / r	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	0.1225	0.1108	0.1003	0.0907	0.0821	0.0743	0.0672	0.0608	0.0550	0.0498
1	0.2572	0.2438	0.2306	0.2177	0.2052	0.1931	0.1815	0.1703	0.1596	0.1494
2	0.2700	0.2681	0.2652	0.2613	0.2565	0.2510	0.2450	0.2384	0.2314	0.2240
3	0.1890	0.1966	0.2033	0.2090	0.2138	0.2176	0.2205	0.2225	0.2237	0.2240
4	0.0992	0.1082	0.1169	0.1254	0.1336	0.1414	0.1488	0.1557	0.1622	0.1680
5	0.0417	0.0476	0.0538	0.0602	0.0668	0.0735	0.0804	0.0872	0.0940	0.1008
6	0.0146	0.0174	0.0206	0.0241	0.0278	0.0319	0.0362	0.0407	0.0455	0.0504
7	0.0044	0.0055	0.0068	0.0083	0.0099	0.0118	0.0139	0.0163	0.0188	0.0216
8	0.0011	0.0015	0.0019	0.0025	0.0031	0.0038	0.0047	0.0057	0.0068	0.0081
9	0.0003	0.0004	0.0005	0.0007	0.0009	0.0011	0.0014	0.0018	0.0022	0.0027
10	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0.0004	0.0005	0.0006	0.0008

续表

r \ μ	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0002	0.0002
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001

r \ μ	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0
0	0.0450	0.0408	0.0369	0.0334	0.0302	0.0273	0.0247	0.0224	0.0202	0.0183
1	0.1397	0.1304	0.1217	0.1135	0.1057	0.0984	0.0915	0.0850	0.0789	0.0733
2	0.2165	0.2087	0.2008	0.1929	0.1850	0.1771	0.1692	0.1615	0.1539	0.1465
3	0.2237	0.2226	0.2209	0.2186	0.2158	0.2125	0.2087	0.2046	0.2001	0.1954
4	0.1733	0.1781	0.1823	0.1858	0.1888	0.1912	0.1931	0.1944	0.1951	0.1954
5	0.1075	0.1140	0.1203	0.1264	0.1322	0.1377	0.1429	0.1477	0.1522	0.1563
6	0.0555	0.0608	0.0662	0.0716	0.0771	0.0826	0.0881	0.0936	0.0989	0.1042
7	0.0246	0.0278	0.0312	0.0348	0.0385	0.0425	0.0466	0.0508	0.0551	0.0595
8	0.0095	0.0111	0.0129	0.0148	0.0169	0.0191	0.0215	0.0241	0.0269	0.0298
9	0.0033	0.0040	0.0047	0.0056	0.0066	0.0076	0.0089	0.0102	0.0116	0.0132
10	0.0010	0.0013	0.0016	0.0019	0.0023	0.0028	0.0033	0.0039	0.0045	0.0053
11	0.0003	0.0004	0.0005	0.0006	0.0007	0.0009	0.0011	0.0013	0.0016	0.0019
12	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0.0003	0.0004	0.0005	0.0006
13	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002
14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001

r \ μ	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	5.0
0	0.0166	0.0150	0.0136	0.0123	0.0111	0.0101	0.0091	0.0082	0.0074	0.0067
1	0.0679	0.0630	0.0583	0.0540	0.0500	0.0462	0.0427	0.0395	0.0365	0.0337
2	0.1393	0.1323	0.1254	0.1188	0.1125	0.1063	0.1005	0.0948	0.0894	0.0842
3	0.1904	0.1852	0.1798	0.1743	0.1687	0.1631	0.1574	0.1517	0.1469	0.1404
4	0.1951	0.1944	0.1932	0.1917	0.1898	0.1875	0.1849	0.1820	0.1789	0.1755
5	0.1600	0.1633	0.1662	0.1687	0.1708	0.1725	0.1738	0.1747	0.1753	0.1755
6	0.1093	0.1143	0.1191	0.1237	0.1281	0.1323	0.1362	0.1398	0.1432	0.1462
7	0.0640	0.0686	0.0732	0.0778	0.0824	0.0869	0.0914	0.0959	0.1002	0.1044
8	0.0328	0.0360	0.0393	0.0428	0.0463	0.0500	0.0537	0.0575	0.0614	0.0653
9	0.0150	0.0168	0.0188	0.0209	0.0232	0.0255	0.0281	0.0307	0.0334	0.0363
10	0.0061	0.0071	0.0081	0.0092	0.0104	0.0118	0.0132	0.0147	0.0614	0.0181
11	0.0023	0.0027	0.0032	0.0037	0.0043	0.0049	0.0056	0.0064	0.0073	0.0082
12	0.0008	0.0009	0.0011	0.0013	0.0016	0.0019	0.0022	0.0026	0.0030	0.0034
13	0.0002	0.0003	0.0004	0.0005	0.0006	0.0007	0.0008	0.0009	0.0011	0.0013
14	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0.0003	0.0004	0.0005
15	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002
16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

续表

r \ μ	5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0
0	0.0061	0.0055	0.0050	0.0045	0.0041	0.0037	0.0033	0.0030	0.0027	0.0025
1	0.0311	0.0287	0.0265	0.0244	0.0225	0.0207	0.0191	0.0176	0.0162	0.0149
2	0.0793	0.0746	0.0701	0.0659	0.0618	0.0580	0.0544	0.0509	0.0477	0.0446
3	0.1348	0.1293	0.1239	0.1185	0.1133	0.1082	0.1033	0.0985	0.0938	0.0892
4	0.1719	0.1681	0.1641	0.1600	0.1558	0.1515	0.1472	0.1428	0.1383	0.1339
5	0.1753	0.1748	0.1740	0.1728	0.1714	0.1697	0.1678	0.1656	0.1632	0.1606
6	0.1490	0.1515	0.1537	0.1555	0.1571	0.1584	0.1594	0.1601	0.1605	0.1606
7	0.1086	0.1125	0.1163	0.1200	0.1234	0.1267	0.1298	0.1326	0.1353	0.1377
8	0.0692	0.0731	0.0771	0.0810	0.0849	0.0887	0.0925	0.0962	0.0998	0.1033
9	0.0392	0.0423	0.0454	0.0486	0.0519	0.0552	0.0586	0.0620	0.0654	0.0688
10	0.0200	0.0220	0.0241	0.0262	0.0285	0.0309	0.0334	0.0359	0.0386	0.0413
11	0.0093	0.0104	0.0116	0.0129	0.0143	0.0157	0.0173	0.0190	0.0207	0.0225
12	0.0039	0.0045	0.0051	0.0058	0.0065	0.0073	0.0082	0.0092	0.0102	0.0113
13	0.0015	0.0018	0.0021	0.0024	0.0028	0.0032	0.0036	0.0041	0.0046	0.0052
14	0.0006	0.0007	0.0008	0.0009	0.0011	0.0013	0.0015	0.0017	0.0019	0.0022
15	0.0002	0.0002	0.0003	0.0003	0.0004	0.0005	0.0006	0.0007	0.0008	0.0009
16	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0003	0.0003
17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001
18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

r \ μ	6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0
0	0.0022	0.0020	0.0018	0.0017	0.0015	0.0014	0.0012	0.0011	0.0010	0.0009
1	0.0137	0.0126	0.0116	0.0106	0.0098	0.0090	0.0082	0.0076	0.0070	0.0064
2	0.0417	0.0390	0.0364	0.0340	0.0318	0.0296	0.0276	0.0258	0.0240	0.0223
3	0.0848	0.0806	0.0765	0.0726	0.0688	0.0652	0.0617	0.0584	0.0552	0.0521
4	0.1294	0.1249	0.1205	0.1162	0.1118	0.1076	0.1034	0.0992	0.0952	0.0912
5	0.1579	0.1549	0.1519	0.1487	0.1454	0.1420	0.1385	0.1349	0.1314	0.1277
6	0.1605	0.1601	0.1595	0.1586	0.1575	0.1562	0.1546	0.1529	0.1511	0.1490
7	0.1399	0.1418	0.1435	0.1450	0.1462	0.1472	0.1480	0.1486	0.1489	0.1490
8	0.1066	0.1099	0.1130	0.1160	0.1188	0.1215	0.1240	0.1263	0.1284	0.1304
9	0.0723	0.0757	0.0791	0.0825	0.0858	0.0891	0.0923	0.0954	0.0985	0.1014
10	0.0441	0.0469	0.0498	0.0528	0.0558	0.0588	0.0618	0.0649	0.0679	0.0710
11	0.0244	0.0265	0.0285	0.0307	0.0330	0.0353	0.0377	0.0401	0.0426	0.0452
12	0.0124	0.0137	0.0150	0.0164	0.0179	0.0194	0.0210	0.0227	0.0245	0.0263
13	0.0058	0.0065	0.0073	0.0081	0.0089	0.0099	0.0108	0.0119	0.0130	0.0142
14	0.0025	0.0029	0.0033	0.0037	0.0041	0.0046	0.0052	0.0058	0.0064	0.0071
15	0.0010	0.0012	0.0014	0.0016	0.0018	0.0020	0.0023	0.0026	0.0029	0.0033
16	0.0004	0.0005	0.0005	0.0006	0.0007	0.0008	0.0010	0.0011	0.0013	0.0014
17	0.0001	0.0002	0.0002	0.0002	0.0003	0.0003	0.0004	0.0004	0.0005	0.0006
18	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002
19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001

续表

μ / r	7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0
0	0.0008	0.0007	0.0007	0.0006	0.0006	0.0005	0.0005	0.0004	0.0004	0.0003
1	0.0059	0.0054	0.0049	0.0045	0.0041	0.0038	0.0035	0.0032	0.0029	0.0027
2	0.0208	0.0194	0.0180	0.0167	0.0156	0.0145	0.0134	0.0125	0.0116	0.0107
3	0.0492	0.0464	0.0438	0.0413	0.0389	0.0366	0.0345	0.0324	0.0305	0.0286
4	0.0874	0.0836	0.0799	0.0764	0.0729	0.0696	0.0663	0.0632	0.0602	0.0573
5	0.1241	0.1204	0.1167	0.1130	0.1094	0.1057	0.1021	0.0986	0.0951	0.0916
6	0.1468	0.1445	0.1420	0.1394	0.1367	0.1339	0.1311	0.1282	0.1252	0.1221
7	0.1489	0.1486	0.1481	0.1474	0.1465	0.1454	0.1442	0.1428	0.1413	0.1396
8	0.1321	0.1337	0.1351	0.1363	0.1373	0.1381	0.1388	0.1392	0.1395	0.1396
9	0.1042	0.1070	0.1096	0.1121	0.1144	0.1167	0.1187	0.1207	0.1224	0.1241
10	0.0740	0.0770	0.0800	0.0829	0.0858	0.0887	0.0914	0.0941	0.0967	0.0993
11	0.0478	0.0504	0.0531	0.0558	0.0585	0.0613	0.0640	0.0667	0.0695	0.0722
12	0.0283	0.0303	0.0323	0.0344	0.0366	0.0388	0.0411	0.0434	0.0457	0.0481
13	0.0154	0.0168	0.0181	0.0196	0.0211	0.0227	0.0243	0.0260	0.0278	0.0296
14	0.0078	0.0086	0.0095	0.0104	0.0113	0.0123	0.0134	0.0145	0.0157	0.0169
15	0.0037	0.0041	0.0046	0.0051	0.0057	0.0062	0.0069	0.0075	0.0083	0.0090
16	0.0016	0.0019	0.0021	0.0024	0.0026	0.0030	0.0033	0.0037	0.0041	0.0045
17	0.0007	0.0008	0.0009	0.0010	0.0012	0.0013	0.0015	0.0017	0.0019	0.0021
18	0.0003	0.0003	0.0004	0.0004	0.0005	0.0006	0.0006	0.0007	0.0008	0.0009
19	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0003	0.0003	0.0003	0.0004
20	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002
21	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001

μ / r	8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0
0	0.0003	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0001	0.0001
1	0.0025	0.0023	0.0021	0.0019	0.0017	0.0016	0.0014	0.0013	0.0012	0.0011
2	0.0100	0.0092	0.0086	0.0079	0.0074	0.0068	0.0063	0.0058	0.0054	0.0050
3	0.0269	0.0252	0.0237	0.0222	0.0208	0.0195	0.0183	0.0171	0.0160	0.0150
4	0.0544	0.0517	0.0491	0.0466	0.0443	0.0420	0.0398	0.0377	0.0357	0.0337
5	0.0882	0.0849	0.0816	0.0784	0.0752	0.0722	0.0692	0.0663	0.0635	0.0607
6	0.1191	0.1160	0.1128	0.1097	0.1066	0.1034	0.1003	0.0972	0.0941	0.0911
7	0.1378	0.1358	0.1338	0.1317	0.1294	0.1271	0.1247	0.1222	0.1197	0.1171
8	0.1395	0.1392	0.1388	0.1382	0.1375	0.1366	0.1356	0.1344	0.1332	0.1318
9	0.1256	0.1269	0.1280	0.1290	0.1299	0.1306	0.1311	0.1315	0.1317	0.1318
10	0.1017	0.1040	0.1063	0.1084	0.1104	0.1123	0.1140	0.1157	0.1172	0.1186
11	0.0749	0.0776	0.0802	0.0828	0.0853	0.0878	0.0902	0.0925	0.0948	0.0970
12	0.0505	0.0530	0.0555	0.0579	0.0604	0.0629	0.0654	0.0679	0.0703	0.0728
13	0.0315	0.0334	0.0354	0.0374	0.0395	0.0416	0.0438	0.0459	0.0481	0.0504
14	0.0182	0.0196	0.0210	0.0225	0.0240	0.0256	0.0272	0.0289	0.0306	0.0324

续表

r \ μ	8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0
15	0.0098	0.0107	0.0116	0.0126	0.0136	0.0147	0.0158	0.0169	0.0182	0.0194
16	0.0050	0.0055	0.0060	0.0066	0.0072	0.0079	0.0086	0.0093	0.0101	0.0109
17	0.0024	0.0026	0.0029	0.0033	0.0036	0.0040	0.0044	0.0048	0.0053	0.0058
18	0.0011	0.0012	0.0014	0.0015	0.0017	0.0019	0.0021	0.0024	0.0026	0.0029
19	0.0005	0.0005	0.0006	0.0007	0.0008	0.0009	0.0010	0.0011	0.0012	0.0014
20	0.0002	0.0002	0.0002	0.0003	0.0003	0.0004	0.0004	0.0005	0.0005	0.0006
21	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002	0.0003
22	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
23	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

r \ μ	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0
0	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000
1	0.0010	0.0009	0.0009	0.0008	0.0007	0.0007	0.0006	0.0005	0.0005	0.0005
2	0.0046	0.0043	0.0040	0.0037	0.0034	0.0031	0.0029	0.0027	0.0025	0.0023
3	0.0140	0.0131	0.0123	0.0115	0.0107	0.0100	0.0093	0.0087	0.0081	0.0076
4	0.0319	0.0302	0.0285	0.0269	0.0254	0.0240	0.0226	0.0213	0.0201	0.0189
5	0.0581	0.0555	0.0530	0.0506	0.0483	0.0460	0.0439	0.0418	0.0398	0.0378
6	0.0881	0.0851	0.0822	0.0793	0.0764	0.0736	0.0709	0.0682	0.0656	0.0631
7	0.1145	0.1118	0.1091	0.1064	0.1037	0.1010	0.0982	0.0955	0.0928	0.0901
8	0.1302	0.1286	0.1269	0.1251	0.1232	0.1212	0.1191	0.1170	0.1148	0.1126
9	0.1317	0.1315	0.1311	0.1306	0.1300	0.1293	0.1284	0.1274	0.1263	0.1251
10	0.1198	0.1210	0.1219	0.1228	0.1235	0.1241	0.1245	0.1249	0.1250	0.1251
11	0.0991	0.1012	0.1031	0.1049	0.1067	0.1083	0.1098	0.1112	0.1125	0.1137
12	0.0752	0.0776	0.0799	0.0822	0.0844	0.0866	0.0888	0.0908	0.0928	0.0948
13	0.0526	0.0549	0.0572	0.0594	0.0617	0.0640	0.0662	0.0685	0.0707	0.0729
14	0.0342	0.0361	0.0380	0.0399	0.0419	0.0439	0.0459	0.0479	0.0500	0.0521
15	0.0208	0.0221	0.0235	0.0250	0.0265	0.0281	0.0297	0.0313	0.0330	0.0347
16	0.0118	0.0127	0.0137	0.0147	0.0157	0.0168	0.0180	0.0192	0.0204	0.0217
17	0.0063	0.0069	0.0075	0.0081	0.0088	0.0095	0.0103	0.0111	0.0119	0.0128
18	0.0032	0.0035	0.0039	0.0042	0.0046	0.0051	0.0055	0.0060	0.0065	0.0071
19	0.0015	0.0017	0.0019	0.0021	0.0023	0.0026	0.0028	0.0031	0.0034	0.0037
20	0.0007	0.0008	0.0009	0.0010	0.0011	0.0012	0.0014	0.0015	0.0017	0.0019
21	0.0003	0.0003	0.0004	0.0004	0.0005	0.0006	0.0006	0.0007	0.0008	0.0009
22	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002	0.0003	0.0003	0.0004	0.0004
23	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002
24	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0001

r \ μ	11.0	12.0	13.0	14.0	15.0	16.0	17.0	18.0	19.0	20.0
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

续表

r \ μ	11.0	12.0	13.0	14.0	15.0	16.0	17.0	18.0	19.0	20.0
2	0.0010	0.0004	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.0037	0.0018	0.0008	0.0004	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000
4	0.0102	0.0053	0.0027	0.0013	0.0006	0.0003	0.0001	0.0001	0.0000	0.0000
5	0.0224	0.0127	0.0070	0.0037	0.0019	0.0010	0.0005	0.0002	0.0001	0.0001
6	0.0411	0.0255	0.0152	0.0087	0.0048	0.0026	0.0014	0.0007	0.0004	0.0002
7	0.0646	0.0437	0.0281	0.0174	0.0104	0.0060	0.0034	0.0019	0.0010	0.0005
8	0.0888	0.0655	0.0457	0.0304	0.0194	0.0120	0.0072	0.0042	0.0024	0.0013
9	0.1085	0.0874	0.0661	0.0473	0.0324	0.0213	0.0135	0.0083	0.0050	0.0029
10	0.1194	0.1048	0.0859	0.0663	0.0486	0.0341	0.0230	0.0150	0.0095	0.0058
11	0.1194	0.1144	0.1015	0.0844	0.0663	0.0496	0.0355	0.0245	0.0164	0.0106
12	0.1094	0.1144	0.1099	0.0984	0.0829	0.0661	0.0504	0.0368	0.0259	0.0176
13	0.0926	0.1056	0.1099	0.1060	0.0956	0.0814	0.0658	0.0509	0.0378	0.0271
14	0.0728	0.0905	0.1021	0.1060	0.1024	0.0930	0.0800	0.0655	0.0514	0.0387
15	0.0534	0.0724	0.0885	0.0989	0.1024	0.0992	0.0906	0.0786	0.0650	0.0516
16	0.0367	0.0543	0.0719	0.0866	0.0960	0.0992	0.0963	0.0884	0.0772	0.0646
17	0.0237	0.0383	0.0550	0.0713	0.0847	0.0934	0.0963	0.0936	0.0863	0.0760
18	0.0145	0.0255	0.0397	0.0554	0.0706	0.0830	0.0909	0.0936	0.0911	0.0844
19	0.0084	0.0161	0.0272	0.0409	0.0557	0.0699	0.0814	0.0887	0.0911	0.0888
20	0.0046	0.0097	0.0177	0.0286	0.0418	0.0559	0.0692	0.0798	0.0866	0.0888
21	0.0024	0.0055	0.0109	0.0191	0.0299	0.0426	0.0560	0.0684	0.0783	0.0846
22	0.0012	0.0030	0.0065	0.0121	0.0204	0.0310	0.0433	0.0560	0.0676	0.0769
23	0.0006	0.0016	0.0037	0.0074	0.0133	0.0216	0.0320	0.0438	0.0559	0.0669
24	0.0003	0.0008	0.0020	0.0043	0.0083	0.0144	0.0226	0.0328	0.0442	0.0557
25	0.0001	0.0004	0.0010	0.0024	0.0050	0.0092	0.0154	0.0237	0.0336	0.0446
26	0.0000	0.0002	0.0005	0.0013	0.0029	0.0057	0.0101	0.0164	0.0246	0.0343
27	0.0000	0.0001	0.0002	0.0007	0.0016	0.0034	0.0063	0.0109	0.0173	0.0254
28	0.0000	0.0000	0.0001	0.0003	0.0009	0.0019	0.0038	0.0070	0.0117	0.0181
29	0.0000	0.0000	0.0001	0.0002	0.0004	0.0011	0.0023	0.0044	0.0077	0.0125
30	0.0000	0.0000	0.0000	0.0001	0.0002	0.0006	0.0013	0.0026	0.0049	0.0083
31	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0007	0.0015	0.0030	0.0054
32	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0004	0.0009	0.0018	0.0034
33	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0005	0.0010	0.0020
34	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0006	0.0012
35	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0007
36	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0004
37	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002
38	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
39	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
40	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

表 4　累积泊松分布表

$$F(x;\ \mu)=\sum_{r=0}^{x} P(r;\ \mu)$$

x \ μ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066	0.3679
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781	0.8442	0.8088	0.7725	0.7358
2	0.9998	0.9989	0.9964	0.9921	0.9856	0.9769	0.9659	0.9526	0.9371	0.9197
3	1.0000	0.9999	0.9997	0.9992	0.9982	0.9966	0.9942	0.9909	0.9865	0.9810
4	1.0000	1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9986	0.9977	0.9963
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9997	0.9994
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

x \ μ	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	0.3329	0.3012	0.2725	0.2466	0.2231	0.2019	0.1827	0.1653	0.1496	0.1353
1	0.6990	0.6626	0.6268	0.5918	0.5578	0.5249	0.4932	0.4628	0.4337	0.4060
2	0.9004	0.8795	0.8571	0.8335	0.8088	0.7834	0.7572	0.7306	0.7037	0.6767
3	0.9743	0.9662	0.9569	0.9463	0.9344	0.9212	0.9068	0.8913	0.8747	0.8571
4	0.9946	0.9923	0.9893	0.9857	0.9814	0.9763	0.9704	0.9636	0.9559	0.9473
5	0.9990	0.9985	0.9978	0.9968	0.9955	0.9940	0.9920	0.9896	0.9868	0.9834
6	0.9999	0.9997	0.9996	0.9994	0.9991	0.9987	0.9981	0.9974	0.9966	0.9955
7	1.0000	1.0000	0.9999	0.9999	0.9998	0.9997	0.9996	0.9994	0.9992	0.9989
8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9998	0.9998
9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

x \ μ	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	0.1225	0.1108	0.1003	0.0907	0.0821	0.0743	0.0672	0.0608	0.0550	0.0498
1	0.3796	0.3546	0.3309	0.3084	0.2873	0.2674	0.2487	0.2311	0.2146	0.1991
2	0.6496	0.6227	0.5960	0.5697	0.5438	0.5184	0.4936	0.4695	0.4460	0.4232
3	0.8386	0.8194	0.7993	0.7787	0.7576	0.7360	0.7141	0.6919	0.6696	0.6472
4	0.9379	0.9275	0.9162	0.9041	0.8912	0.8774	0.8629	0.8477	0.8318	0.8153
5	0.9796	0.9751	0.9700	0.9643	0.9580	0.9510	0.9433	0.9349	0.9258	0.9161
6	0.9941	0.9925	0.9906	0.9884	0.9858	0.9828	0.9794	0.9756	0.9713	0.9665
7	0.9985	0.9980	0.9974	0.9967	0.9958	0.9947	0.9934	0.9919	0.9901	0.9881
8	0.9997	0.9995	0.9994	0.9991	0.9989	0.9985	0.9981	0.9976	0.9969	0.9962
9	0.9999	0.9999	0.9999	0.9998	0.9997	0.9996	0.9995	0.9993	0.9991	0.9989
10	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9999	0.9998	0.9998	0.9997

续表

x \ μ	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

x \ μ	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0
0	0.0450	0.0408	0.0369	0.0334	0.0302	0.0273	0.0247	0.0224	0.0202	0.0183
1	0.1847	0.1712	0.1586	0.1468	0.1359	0.1257	0.1162	0.1074	0.0992	0.0916
2	0.4012	0.3799	0.3594	0.3397	0.3208	0.3027	0.2854	0.2689	0.2531	0.2381
3	0.6248	0.6025	0.5803	0.5584	0.5366	0.5152	0.4942	0.4735	0.4532	0.4335
4	0.7982	0.7806	0.7626	0.7442	0.7254	0.7064	0.6872	0.6678	0.6484	0.6288
5	0.9057	0.8946	0.8829	0.8705	0.8576	0.8441	0.8301	0.8156	0.8006	0.7851
6	0.9612	0.9554	0.9490	0.9421	0.9347	0.9267	0.9182	0.9091	0.8995	0.8893
7	0.9858	0.9832	0.9802	0.9769	0.9733	0.9692	0.9648	0.9599	0.9546	0.9489
8	0.9953	0.9943	0.9931	0.9917	0.9901	0.9883	0.9863	0.9840	0.9815	0.9786
9	0.9986	0.9982	0.9978	0.9973	0.9967	0.9960	0.9952	0.9942	0.9931	0.9919
10	0.9996	0.9995	0.9994	0.9992	0.9990	0.9987	0.9984	0.9981	0.9977	0.9972
11	0.9999	0.9999	0.9998	0.9998	0.9997	0.9996	0.9995	0.9994	0.9993	0.9991
12	1.0000	1.0000	1.0000	0.9999	0.9999	0.9999	0.9999	0.9998	0.9998	0.9997
13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999
14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

x \ μ	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	5.0
0	0.0166	0.0150	0.0136	0.0123	0.0111	0.0101	0.0091	0.0082	0.0074	0.0067
1	0.0845	0.0780	0.0719	0.0663	0.0611	0.0563	0.0518	0.0477	0.0439	0.0404
2	0.2238	0.2102	0.1974	0.1851	0.1736	0.1626	0.1523	0.1425	0.1333	0.1247
3	0.4142	0.3954	0.3772	0.3594	0.3423	0.3257	0.3097	0.2942	0.2793	0.2650
4	0.6093	0.5898	0.5704	0.5512	0.5321	0.5132	0.4946	0.4763	0.4582	0.4405
5	0.7693	0.7531	0.7367	0.7199	0.7029	0.6858	0.6684	0.6510	0.6335	0.6160
6	0.8786	0.8675	0.8558	0.8436	0.8311	0.8180	0.8046	0.7908	0.7767	0.7622
7	0.9427	0.9361	0.9290	0.9214	0.9134	0.9049	0.8960	0.8867	0.8769	0.8666
8	0.9755	0.9721	0.9683	0.9642	0.9597	0.9549	0.9497	0.9442	0.9382	0.9319
9	0.9905	0.9889	0.9871	0.9851	0.9829	0.9805	0.9778	0.9749	0.9717	0.9682
10	0.9966	0.9959	0.9952	0.9943	0.9933	0.9922	0.9910	0.9896	0.9880	0.9863
11	0.9989	0.9986	0.9983	0.9980	0.9976	0.9971	0.9966	0.9960	0.9953	0.9945
12	0.9997	0.9996	0.9995	0.9993	0.9992	0.9990	0.9988	0.9986	0.9983	0.9980
13	0.9999	0.9999	0.9998	0.9998	0.9997	0.9997	0.9996	0.9995	0.9994	0.9993
14	1.0000	1.0000	1.0000	0.9999	0.9999	0.9999	0.9999	0.9999	0.9998	0.9998
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999
16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

续表

x \ μ	5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0
0	0.0061	0.0055	0.0050	0.0045	0.0041	0.0037	0.0033	0.0030	0.0027	0.0025
1	0.0372	0.0342	0.0314	0.0289	0.0266	0.0244	0.0224	0.0206	0.0189	0.0174
2	0.1165	0.1088	0.1016	0.0948	0.0884	0.0824	0.0768	0.0715	0.0666	0.0620
3	0.2513	0.2381	0.2254	0.2133	0.2017	0.1906	0.1800	0.1700	0.1604	0.1512
4	0.4231	0.4061	0.3895	0.3733	0.3575	0.3422	0.3272	0.3127	0.2987	0.2851
5	0.5984	0.5809	0.5635	0.5461	0.5289	0.5119	0.4950	0.4783	0.4619	0.4457
6	0.7474	0.7324	0.7171	0.7017	0.6860	0.6703	0.6544	0.6384	0.6224	0.6063
7	0.8560	0.8449	0.8335	0.8217	0.8095	0.7970	0.7841	0.7710	0.7576	0.7440
8	0.9252	0.9181	0.9106	0.9027	0.8944	0.8857	0.8766	0.8672	0.8574	0.8472
9	0.9644	0.9603	0.9559	0.9512	0.9462	0.9409	0.9352	0.9292	0.9228	0.9161
10	0.9844	0.9823	0.9800	0.9775	0.9747	0.9718	0.9686	0.9651	0.9614	0.9574
11	0.9937	0.9927	0.9916	0.9904	0.9890	0.9875	0.9859	0.9841	0.9821	0.9799
12	0.9976	0.9972	0.9967	0.9962	0.9955	0.9949	0.9941	0.9932	0.9922	0.9912
13	0.9992	0.9990	0.9988	0.9986	0.9983	0.9980	0.9977	0.9973	0.9969	0.9964
14	0.9997	0.9997	0.9996	0.9995	0.9994	0.9993	0.9991	0.9990	0.9988	0.9986
15	0.9999	0.9999	0.9999	0.9998	0.9998	0.9998	0.9997	0.9996	0.9996	0.9995
16	1.0000	1.0000	1.0000	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9998
17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

x \ μ	6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0
0	0.0022	0.0020	0.0018	0.0017	0.0015	0.0014	0.0012	0.0011	0.0010	0.0009
1	0.0159	0.0146	0.0134	0.0123	0.0113	0.0103	0.0095	0.0087	0.0080	0.0073
2	0.0577	0.0536	0.0498	0.0463	0.0430	0.0400	0.0371	0.0344	0.0320	0.0296
3	0.1425	0.1342	0.1264	0.1189	0.1118	0.1052	0.0988	0.0928	0.0871	0.0818
4	0.2719	0.2592	0.2469	0.2351	0.2237	0.2127	0.2022	0.1920	0.1823	0.1730
5	0.4298	0.4141	0.3988	0.3837	0.3690	0.3547	0.3406	0.3270	0.3137	0.3007
6	0.5902	0.5742	0.5582	0.5423	0.5265	0.5108	0.4953	0.4799	0.4647	0.4497
7	0.7301	0.7160	0.7017	0.6873	0.6728	0.6581	0.6433	0.6285	0.6136	0.5987
8	0.8367	0.8259	0.8148	0.8033	0.7916	0.7796	0.7673	0.7548	0.7420	0.7291
9	0.9090	0.9016	0.8939	0.8858	0.8774	0.8686	0.8596	0.8502	0.8405	0.8305
10	0.9531	0.9486	0.9437	0.9386	0.9332	0.9274	0.9214	0.9151	0.9084	0.9015
11	0.9776	0.9750	0.9723	0.9693	0.9661	0.9627	0.9591	0.9552	0.9510	0.9467
12	0.9900	0.9887	0.9873	0.9857	0.9840	0.9821	0.9801	0.9779	0.9755	0.9730
13	0.9958	0.9952	0.9945	0.9937	0.9929	0.9920	0.9909	0.9898	0.9885	0.9872
14	0.9984	0.9981	0.9978	0.9974	0.9970	0.9966	0.9961	0.9956	0.9950	0.9943
15	0.9994	0.9993	0.9992	0.9990	0.9988	0.9986	0.9984	0.9982	0.9979	0.9976
16	0.9998	0.9997	0.9997	0.9996	0.9996	0.9995	0.9994	0.9993	0.9992	0.9990
17	0.9999	0.9999	0.9999	0.9999	0.9998	0.9998	0.9998	0.9997	0.9997	0.9996
18	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

续表

x \ μ	7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0
0	0.0008	0.0007	0.0007	0.0006	0.0006	0.0005	0.0005	0.0004	0.0004	0.0003
1	0.0067	0.0061	0.0056	0.0051	0.0047	0.0043	0.0039	0.0036	0.0033	0.0020
2	0.0275	0.0255	0.0236	0.0219	0.0203	0.0188	0.0174	0.0161	0.0149	0.0138
3	0.0767	0.0719	0.0674	0.0632	0.0591	0.0554	0.0518	0.0485	0.0453	0.0424
4	0.1641	0.1555	0.1473	0.1395	0.1321	0.1249	0.1181	0.1117	0.1055	0.0996
5	0.2881	0.2759	0.2640	0.2526	0.2414	0.2307	0.2203	0.2103	0.2006	0.1912
6	0.4349	0.4204	0.4060	0.3920	0.3782	0.3646	0.3514	0.3384	0.3257	0.3134
7	0.5838	0.5689	0.5541	0.5393	0.5246	0.5100	0.4956	0.4812	0.4670	0.4530
8	0.7160	0.7027	0.6892	0.6757	0.6620	0.6482	0.6343	0.6204	0.6065	0.5925
9	0.8202	0.8096	0.7988	0.7877	0.7764	0.7649	0.7531	0.7411	0.7290	0.7166
10	0.8942	0.8867	0.8788	0.8707	0.8622	0.8535	0.8445	0.8352	0.8257	0.8159
11	0.9420	0.9371	0.9319	0.9265	0.9208	0.9148	0.9085	0.9020	0.8952	0.8881
12	0.9703	0.9673	0.9642	0.9609	0.9573	0.9536	0.9496	0.9454	0.9409	0.9362
13	0.9857	0.9841	0.9824	0.9805	0.9784	0.9762	0.9739	0.9714	0.9687	0.9658
14	0.9935	0.9927	0.9918	0.9908	0.9897	0.9886	0.9873	0.9859	0.9844	0.9827
15	0.9972	0.9969	0.9964	0.9959	0.9954	0.9948	0.9941	0.9934	0.9926	0.9918
16	0.9989	0.9987	0.9985	0.9983	0.9980	0.9978	0.9974	0.9971	0.9967	0.9963
17	0.9996	0.9995	0.9994	0.9993	0.9992	0.9991	0.9989	0.9988	0.9986	0.9984
18	0.9998	0.9998	0.9998	0.9997	0.9997	0.9996	0.9996	0.9995	0.9994	0.9993
19	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9998	0.9998	0.9998	0.9997
20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9999	0.9999
21	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

x \ μ	8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0
0	0.0003	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0001	0.0001
1	0.0028	0.0025	0.0023	0.0021	0.0019	0.0018	0.0016	0.0015	0.0014	0.0012
2	0.0127	0.0118	0.0109	0.0100	0.0093	0.0086	0.0079	0.0073	0.0068	0.0062
3	0.0396	0.0370	0.0346	0.0323	0.0301	0.0281	0.0262	0.0244	0.0228	0.0212
4	0.0940	0.0887	0.0837	0.0789	0.0744	0.0701	0.0660	0.0621	0.0584	0.0550
5	0.1822	0.1736	0.1653	0.1573	0.1496	0.1422	0.1352	0.1284	0.1219	0.1157
6	0.3013	0.2896	0.2781	0.2670	0.2562	0.2457	0.2355	0.2256	0.2160	0.2068
7	0.4391	0.4254	0.4119	0.3987	0.3856	0.3728	0.3602	0.3478	0.3357	0.3239
8	0.5786	0.5647	0.5507	0.5369	0.5231	0.5094	0.4958	0.4823	0.4689	0.4557
9	0.7041	0.6915	0.6788	0.6659	0.6530	0.6400	0.6269	0.6137	0.6006	0.5874
10	0.8058	0.7955	0.7850	0.7743	0.7634	0.7522	0.7409	0.7294	0.7178	0.7060
11	0.8807	0.8731	0.8652	0.8571	0.8487	0.8400	0.8311	0.8220	0.8126	0.8030
12	0.9313	0.9261	0.9207	0.9150	0.9091	0.9029	0.8965	0.8898	0.8829	0.8758
13	0.9628	0.9595	0.9561	0.9524	0.9486	0.9445	0.9403	0.9358	0.9311	0.9261
14	0.9810	0.9791	0.9771	0.9749	0.9726	0.9701	0.9675	0.9647	0.9617	0.9585

续表

x \ μ	8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0
15	0.9908	0.9898	0.9887	0.9875	0.9862	0.9848	0.9832	0.9816	0.9798	0.9780
16	0.9958	0.9953	0.9947	0.9941	0.9934	0.9926	0.9918	0.9909	0.9899	0.9889
17	0.9982	0.9979	0.9977	0.9973	0.9970	0.9966	0.9962	0.9957	0.9952	0.9947
18	0.9992	0.9991	0.9990	0.9989	0.9987	0.9985	0.9983	0.9981	0.9978	0.9976
19	0.9997	0.9997	0.9996	0.9995	0.9995	0.9994	0.9993	0.9992	0.9991	0.9989
20	0.9999	0.9999	0.9998	0.9998	0.9998	0.9998	0.9997	0.9997	0.9996	0.9996
21	1.0000	1.0000	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9998	0.9998
22	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999
23	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

x \ μ	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0
0	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000
1	0.0011	0.0010	0.0009	0.0009	0.0008	0.0007	0.0007	0.0006	0.0005	0.0005
2	0.0058	0.0053	0.0049	0.0045	0.0042	0.0038	0.0035	0.0033	0.0030	0.0028
3	0.0198	0.0184	0.0172	0.0160	0.0149	0.0138	0.0129	0.0120	0.0111	0.0103
4	0.0517	0.0486	0.0456	0.0429	0.0403	0.0378	0.0355	0.0333	0.0312	0.0293
5	0.1098	0.1041	0.0986	0.0935	0.0885	0.0838	0.0793	0.0750	0.0710	0.0671
6	0.1978	0.1892	0.1808	0.1727	0.1649	0.1574	0.1502	0.1433	0.1366	0.1301
7	0.3123	0.3010	0.2900	0.2792	0.2687	0.2584	0.2485	0.2388	0.2294	0.2202
8	0.4426	0.4296	0.4168	0.4042	0.3918	0.3796	0.3676	0.3558	0.3442	0.3328
9	0.5742	0.5611	0.5479	0.5349	0.5218	0.5089	0.4960	0.4832	0.4705	0.4579
10	0.6941	0.6820	0.6699	0.6576	0.6453	0.6329	0.6205	0.6080	0.5955	0.5830
11	0.7932	0.7832	0.7730	0.7626	0.7520	0.7412	0.7303	0.7193	0.7081	0.6968
12	0.8684	0.8607	0.8529	0.8448	0.8364	0.8279	0.8191	0.8101	0.8009	0.7916
13	0.9210	0.9156	0.9100	0.9042	0.8981	0.8919	0.8853	0.8786	0.8716	0.8645
14	0.9552	0.9517	0.9480	0.9441	0.9400	0.9357	0.9312	0.9265	0.9216	0.9165
15	0.9760	0.9738	0.9715	0.9691	0.9665	0.9638	0.9609	0.9579	0.9546	0.9513
16	0.9878	0.9865	0.9852	0.9838	0.9823	0.9806	0.9789	0.9770	0.9751	0.9730
17	0.9941	0.9934	0.9927	0.9919	0.9911	0.9902	0.9892	0.9881	0.9870	0.9857
18	0.9973	0.9969	0.9966	0.9962	0.9957	0.9952	0.9947	0.9941	0.9935	0.9928
19	0.9988	0.9986	0.9985	0.9983	0.9980	0.9978	0.9975	0.9972	0.9969	0.9965
20	0.9995	0.9994	0.9993	0.9992	0.9991	0.9990	0.9989	0.9987	0.9986	0.9984
21	0.9998	0.9998	0.9997	0.9997	0.9996	0.9996	0.9995	0.9995	0.9994	0.9993
22	0.9999	0.9999	0.9999	0.9999	0.9999	0.9998	0.9998	0.9998	0.9997	0.9997
23	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
24	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

x \ μ	11.0	12.0	13.0	14.0	15.0	16.0	17.0	18.0	19.0	20.0
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

续表

x \ μ	11.0	12.0	13.0	14.0	15.0	16.0	17.0	18.0	19.0	20.0
2	0.0012	0.0005	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.0049	0.0023	0.0011	0.0005	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000
4	0.0151	0.0076	0.0037	0.0018	0.0009	0.0004	0.0002	0.0001	0.0000	0.0000
5	0.0375	0.0203	0.0107	0.0055	0.0028	0.0014	0.0007	0.0003	0.0002	0.0001
6	0.0786	0.0458	0.0259	0.0142	0.0076	0.0040	0.0021	0.0010	0.0005	0.0003
7	0.1432	0.0895	0.0540	0.0316	0.0180	0.0100	0.0054	0.0029	0.0015	0.0008
8	0.2320	0.1550	0.0998	0.0621	0.0374	0.0220	0.0126	0.0071	0.0039	0.0021
9	0.3405	0.2424	0.1658	0.1094	0.0699	0.0433	0.0261	0.0154	0.0089	0.0050
10	0.4599	0.3472	0.2517	0.1757	0.1185	0.0774	0.0491	0.0304	0.0183	0.0108
11	0.5793	0.4616	0.3532	0.2600	0.1848	0.1270	0.0847	0.0549	0.0347	0.0214
12	0.6887	0.5760	0.4631	0.3585	0.2676	0.1931	0.1350	0.0917	0.0606	0.0390
13	0.7813	0.6815	0.5730	0.4644	0.3632	0.2745	0.2009	0.1426	0.0984	0.0661
14	0.8540	0.7720	0.6751	0.5704	0.4657	0.3675	0.2808	0.2081	0.1497	0.1049
15	0.9074	0.8444	0.7636	0.6694	0.5681	0.4667	0.3715	0.2867	0.2148	0.1565
16	0.9441	0.8987	0.8355	0.7559	0.6641	0.5660	0.4677	0.3751	0.2920	0.2211
17	0.9678	0.9370	0.8905	0.8272	0.7489	0.6593	0.5640	0.4686	0.3784	0.2970
18	0.9823	0.9626	0.9302	0.8826	0.8195	0.7423	0.6550	0.5622	0.4695	0.3814
19	0.9907	0.9787	0.9573	0.9235	0.8752	0.8122	0.7363	0.6509	0.5606	0.4703
20	0.9953	0.9884	0.9750	0.9521	0.9170	0.8682	0.8055	0.7307	0.6472	0.5591
21	0.9977	0.9939	0.9859	0.9712	0.9469	0.9108	0.8615	0.7991	0.7255	0.6437
22	0.9990	0.9970	0.9924	0.9833	0.9673	0.9418	0.9047	0.8551	0.7931	0.7206
23	0.9995	0.9985	0.9960	0.9907	0.9805	0.9633	0.9367	0.8989	0.8490	0.7875
24	0.9998	0.9993	0.9980	0.9950	0.9888	0.9777	0.9594	0.9317	0.8933	0.8432
25	0.9999	0.9997	0.9990	0.9974	0.9938	0.9869	0.9748	0.9554	0.9269	0.8878
26	1.0000	0.9999	0.9995	0.9987	0.9967	0.9925	0.9848	0.9718	0.9514	0.9221
27	1.0000	0.9999	0.9998	0.9994	0.9983	0.9959	0.9912	0.9827	0.9687	0.9475
28	1.0000	1.0000	0.9999	0.9997	0.9991	0.9978	0.9950	0.9897	0.9805	0.9657
29	1.0000	1.0000	1.0000	0.9999	0.9996	0.9989	0.9973	0.9941	0.9882	0.9782
30	1.0000	1.0000	1.0000	0.9999	0.9998	0.9994	0.9986	0.9967	0.9930	0.9865
31	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9993	0.9982	0.9960	0.9919
32	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9990	0.9978	0.9953
33	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9995	0.9988	0.9973
34	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9994	0.9985
35	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9992
36	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9996
37	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998
38	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
39	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
40	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

表 5　标准正态分布概率密度表

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad 0 \leqslant x \leqslant 4.99$$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.39894	0.39892	0.39886	0.39876	0.39862	0.39844	0.39822	0.39797	0.39767	0.39733
0.1	0.39695	0.39654	0.39608	0.39559	0.39505	0.39448	0.39387	0.39322	0.39253	0.39181
0.2	0.39104	0.39024	0.38940	0.38853	0.38762	0.38667	0.38568	0.38466	0.38361	0.38251
0.3	0.38139	0.38023	0.37903	0.37780	0.37654	0.37524	0.37391	0.37255	0.37115	0.36973
0.4	0.36827	0.36678	0.36526	0.36371	0.36213	0.36053	0.35889	0.35723	0.35553	0.35381
0.5	0.35207	0.35029	0.34849	0.34667	0.34482	0.34294	0.34105	0.33912	0.33718	0.33521
0.6	0.33322	0.33121	0.32918	0.32713	0.32506	0.32297	0.32086	0.31874	0.31659	0.31443
0.7	0.31225	0.31006	0.30785	0.30563	0.30339	0.30114	0.29887	0.29659	0.29431	0.29200
0.8	0.28969	0.28737	0.28504	0.28269	0.28034	0.27798	0.27562	0.27324	0.27086	0.26848
0.9	0.26609	0.26369	0.26129	0.25888	0.25647	0.25406	0.25164	0.24923	0.24681	0.24439
1.0	0.24197	0.23955	0.23713	0.23471	0.23230	0.22988	0.22747	0.22506	0.22265	0.22025
1.1	0.21785	0.21546	0.21307	0.21069	0.20831	0.20594	0.20357	0.20121	0.19886	0.19652
1.2	0.19419	0.19186	0.18954	0.18724	0.18494	0.18265	0.18037	0.17810	0.17585	0.17360
1.3	0.17137	0.16915	0.16694	0.16474	0.16256	0.16038	0.15822	0.15608	0.15395	0.15183
1.4	0.14973	0.14764	0.14556	0.14350	0.14146	0.13943	0.13742	0.13542	0.13344	0.13147
1.5	0.12952	0.12758	0.12566	0.12376	0.12188	0.12001	0.11816	0.11632	0.11450	0.11270
1.6	0.11092	0.10915	0.10741	0.10567	0.10396	0.10226	0.10059	0.09893	0.09728	0.09566
1.7	0.09405	0.09246	0.09089	0.08933	0.08780	0.08628	0.08478	0.08329	0.08183	0.08038
1.8	0.07895	0.07754	0.07614	0.07477	0.07341	0.07206	0.07074	0.06943	0.06814	0.06687
1.9	0.06562	0.06438	0.06316	0.06195	0.06077	0.05959	0.05844	0.05730	0.05618	0.05508
2.0	0.05399	0.05292	0.05186	0.05082	0.04980	0.04879	0.04780	0.04682	0.04586	0.04491
2.1	0.04398	0.04307	0.04217	0.04128	0.04041	0.03955	0.03871	0.03788	0.03706	0.03626
2.2	0.03547	0.03470	0.03394	0.03319	0.03246	0.03174	0.03103	0.03034	0.02965	0.02898
2.3	0.02833	0.02768	0.02705	0.02643	0.02582	0.02522	0.02463	0.02406	0.02349	0.02294
2.4	0.02239	0.02186	0.02134	0.02083	0.02033	0.01984	0.01936	0.01888	0.01842	0.01797
2.5	0.01753	0.01709	0.01667	0.01625	0.01585	0.01545	0.01506	0.01468	0.01431	0.01394
2.6	0.01358	0.01323	0.01289	0.01256	0.01223	0.01191	0.01160	0.01130	0.01100	0.01071
2.7	0.01042	0.01014	0.00987	0.00961	0.00935	0.00909	0.00885	0.00861	0.00837	0.00814
2.8	0.00792	0.00770	0.00748	0.00727	0.00707	0.00687	0.00668	0.00649	0.00631	0.00613
2.9	0.00595	0.00578	0.00562	0.00545	0.00530	0.00514	0.00499	0.00485	0.00470	0.00457
3.0	0.00443	0.00430	0.00417	0.00405	0.00393	0.00381	0.00370	0.00358	0.00348	0.00337
3.1	0.00327	0.00317	0.00307	0.00298	0.00288	0.00279	0.00271	0.00262	0.00254	0.00246

续表

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
3.2	0.00238	0.00231	0.00224	0.00216	0.00210	0.00203	0.00196	0.00190	0.00184	0.00178
3.3	0.00172	0.00167	0.00161	0.00156	0.00151	0.00146	0.00141	0.00136	0.00132	0.00127
3.4	0.00123	0.00119	0.00115	0.00111	0.00107	0.00104	0.00100	0.00097	0.00094	0.00090
3.5	0.00087	0.00084	0.00081	0.00079	0.00076	0.00073	0.00071	0.00068	0.00066	0.00063
3.6	0.00061	0.00059	0.00057	0.00055	0.00053	0.00051	0.00049	0.00047	0.00046	0.00044
3.7	0.00042	0.00041	0.00039	0.00038	0.00037	0.00035	0.00034	0.00033	0.00031	0.00030
3.8	0.00029	0.00028	0.00027	0.00026	0.00025	0.00024	0.00023	0.00022	0.00021	0.00021
3.9	0.00020	0.00019	0.00018	0.00018	0.00017	0.00016	0.00016	0.00015	0.00014	0.00014
4.0	0.00013	0.00013	0.00012	0.00012	0.00011	0.00011	0.00011	0.00010	0.00010	0.00009
4.1	0.00009	0.00009	0.00008	0.00008	0.00008	0.00007	0.00007	0.00007	0.00006	0.00006
4.2	0.00006	0.00006	0.00005	0.00005	0.00005	0.00005	0.00005	0.00004	0.00004	0.00004
4.3	0.00004	0.00004	0.00004	0.00003	0.00003	0.00003	0.00003	0.00003	0.00003	0.00003
4.4	0.00002	0.00002	0.00002	0.00002	0.00002	0.00002	0.00002	0.00002	0.00002	0.00002
4.5	0.00002	0.00002	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
4.6	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
4.7	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000
4.8	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
4.9	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

表 6 标准正态分布累积分布函数表

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t, \quad 0 \leqslant x \leqslant 4.99. \quad \Phi(-x) = 1 - \Phi(x).$$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147

续表

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	099934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997
4.0	0.99997	0.99997	0.99997	0.99997	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998
4.1	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.99999	0.99999
4.2	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
4.3	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999
4.4	0.99999	0.99999	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
4.5	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
4.6	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
4.7	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
4.8	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
4.9	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

表 7　χ^2 分布的上侧 α 分位数 χ^2_α 表

$$F(\chi^2_\alpha;\nu)=\int_0^{\chi^2_\alpha(v)} f(y;\nu)\mathrm{d}y=1-\alpha,$$

其中 $f(y;\nu)$ 为自由度 v 的 χ^2 分布的概率密度.

ν \ F	0.005	0.010	0.025	0.050	0.100	0.200	0.250	0.500
1	0.00004	0.00016	0.00098	0.00393	0.0158	0.064	0.102	0.455
2	0.0100	0.0201	0.0506	0.103	0.211	0.446	0.575	1.386
3	0.0717	0.115	0.216	0.352	0.584	1.005	1.213	2.366
4	0.207	0.297	0.484	0.711	1.064	1.649	1.923	3.357
5	0.412	0.554	0.831	1.145	1.610	2.343	2.675	4.351
6	0.676	0.872	1.237	1.635	2.204	3.070	3.455	5.348
7	0.989	1.239	1.690	2.167	2.833	3.822	4.255	6.346
8	1.344	1.646	2.180	2.733	3.490	4.594	5.071	7.344
9	1.735	2.088	2.700	3.325	4.168	5.380	5.899	8.343
10	2.156	2.558	3.247	3.940	4.865	6.179	6.737	9.342
11	2.60	3.05	3.82	4.57	5.58	6.99	7.58	10.34
12	3.07	3.57	4.40	5.23	6.30	7.81	8.44	11.34
13	3.57	4.11	5.01	5.89	7.04	8.63	9.30	12.34
14	4.07	4.66	5.63	6.57	7.79	9.47	10.17	13.34
15	4.60	5.23	6.26	7.26	8.55	10.31	11.04	14.34
16	5.14	5.81	6.91	7.96	9.31	11.15	11.91	15.34
17	5.70	6.41	7.56	8.67	10.08	12.00	12.79	16.34
18	6.26	7.01	8.23	9.39	10.86	12.86	13.68	17.34
19	6.84	7.63	8.91	10.12	11.65	13.72	14.56	18.34
20	7.43	8.26	9.59	10.85	12.44	14.58	15.45	19.34
21	8.03	8.90	10.28	11.59	13.24	15.44	16.34	20.34
22	8.64	9.54	10.98	12.34	14.04	16.31	17.24	21.34
23	9.26	10.20	11.69	13.09	14.85	17.19	18.14	22.34
24	9.89	10.86	12.40	13.85	15.66	18.06	19.04	23.34
25	10.52	11.52	13.12	14.61	16.47	18.94	19.94	24.34
26	11.16	12.20	13.84	15.38	17.29	19.82	20.84	25.34
27	11.81	12.88	14.57	16.15	18.11	20.70	21.75	26.34
28	12.46	13.56	15.31	16.93	18.94	21.59	22.66	27.34
29	13.12	14.26	16.05	17.71	19.77	22.48	23.57	28.34

续表

ν \ F	0.005	0.010	0.025	0.050	0.100	0.200	0.250	0.500
30	13.79	14.95	16.79	18.49	20.60	23.36	24.48	29.34
40	20.70	22.16	24.43	26.51	29.05	32.34	33.66	39.34
50	27.99	29.71	32.36	34.76	37.69	41.45	42.94	49.34
60	35.53	37.48	40.48	43.19	46.46	50.64	52.29	59.33
70	43.27	45.44	48.76	51.74	55.33	59.90	61.70	69.33
80	51.17	53.54	57.15	60.39	64.28	69.21	71.14	79.33
90	59.19	61.75	65.64	69.12	73.29	78.56	80.62	89.33
100	67.32	70.06	74.22	77.93	82.36	87.94	90.13	99.33

ν \ F	0.750	0.800	0.900	0.950	0.975	0.990	0.995
1	1.323	1.642	2.706	3.841	5.024	6.635	7.879
2	2.773	3.219	4.605	5.991	7.378	9.210	10.597
3	4.108	4.642	6.251	7.815	9.348	11.345	12.838
4	5.385	5.989	7.779	9.488	11.143	13.277	14.860
5	6.626	7.289	9.236	11.071	12.833	15.086	16.750
6	7.841	8.558	10.645	12.592	14.449	16.812	18.548
7	9.037	9.803	12.017	14.067	16.013	18.475	20.278
8	10.219	11.030	13.362	15.507	17.535	20.090	21.955
9	11.389	12.242	14.684	16.919	19.023	21.666	23.589
10	12.549	13.442	15.987	18.307	20.483	23.209	25.188
11	13.70	14.63	17.28	19.68	21.92	24.73	26.76
12	14.85	15.81	18.55	21.03	23.34	26.22	28.30
13	15.98	16.98	19.81	22.36	24.74	27.69	29.82
14	17.12	18.15	21.06	23.68	26.12	29.14	31.32
15	18.25	19.31	22.31	25.00	27.49	30.58	32.80
16	19.37	20.47	23.54	26.30	28.85	32.00	34.27
17	20.49	21.61	24.77	27.59	30.19	33.41	35.72
18	21.60	22.76	25.99	28.87	31.53	34.81	37.16
19	22.72	23.90	27.20	30.14	32.85	36.19	38.58
20	23.83	25.04	28.41	31.41	34.17	37.57	40.00
21	24.93	26.17	29.62	32.67	35.48	38.93	41.40
22	26.04	27.30	30.81	33.92	36.78	40.29	42.80
23	27.14	28.43	32.01	35.17	38.08	41.64	44.18
24	28.24	29.55	33.20	36.42	39.36	42.98	45.56

续表

ν \ F	0.750	0.800	0.900	0.950	0.975	0.990	0.995
25	29.34	30.68	34.38	37.65	40.65	44.31	46.93
26	30.43	31.79	35.56	38.89	41.92	45.64	48.29
27	31.53	32.91	36.74	40.11	43.19	46.96	49.65
28	32.62	34.03	37.92	41.34	44.46	48.28	50.99
29	33.71	35.14	39.09	42.56	45.72	49.59	52.34
30	34.80	36.25	40.26	43.77	46.98	50.89	53.67
40	45.62	47.27	51.81	55.76	59.34	63.69	66.77
50	56.33	58.16	63.17	67.50	71.42	76.16	79.49
60	66.98	68.97	74.40	79.08	83.30	88.38	91.95
70	77.58	79.72	85.53	90.53	95.03	100.43	104.22
80	88.13	90.41	96.58	101.88	106.63	112.33	116.32
90	98.65	101.06	107.57	113.15	118.14	124.12	128.30
100	109.14	111.67	118.50	124.34	129.56	135.81	140.17

表 8　t 分布的上侧 α 分位数 t_α 表

$$F(t_\alpha;\nu)=\int_{-\infty}^{t_\alpha(\nu)} f(y;\nu)\mathrm{d}y=1-\alpha,$$

其中 $f(y;\nu)$ 为自由度 ν 的 t 分布的概率密度.

$$F(-t;\nu)=1-F(t;\nu).$$

$\nu=\infty$ 相应于标准正态分布.

ν \ F	0.60	0.70	0.80	0.90	0.95	0.975	0.990	0.995	0.999	0.9995
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657	318.31	663.62
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925	22.327	31.598
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169	4.144	4.587

续表

ν \ F	0.60	0.70	0.80	0.90	0.95	0.975	0.990	0.995	0.999	0.9995
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	0.255	0.528	0.849	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390	2.660	3.232	3.460
70	0.254	0.527	0.847	1.294	1.667	1.994	2.381	2.648	3.211	3.435
80	0.254	0.527	0.846	1.292	1.664	1.990	2.374	2.639	3.195	3.416
90	0.254	0.526	0.846	1.291	1.662	1.987	2.369	2.632	3.183	3.402
100	0.254	0.526	0.845	1.290	1.660	1.984	2.364	2.626	3.174	3.391
110	0.254	0.526	0.845	1.289	1.659	1.982	2.361	2.621	3.166	3.381
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576	3.090	3.291

表 9 F 分布的上侧 α 分位数 f_α 表

$$F(f_\alpha;\nu_1,\nu_2)=\int_0^{f_\alpha} f(y;\nu_1,\nu_2)\mathrm{d}y=1-\alpha,$$

其中 $f(y;\nu_1,\nu_2)$ 为自由度 (ν_1,ν_2) 的 F 分布的概率密度.

表中只列出 $F=0.9\sim 0.999$ 的 $f_\alpha(\nu_1,\nu_2)$ 值, $F=0.001\sim 0.1$ 的 $f_\alpha(\nu_1,\nu_2)$ 值可由下面的关系式求出:

$$f_{1-\alpha}(\nu_1,\nu_2)=\frac{1}{f_\alpha(\nu_2,\nu_1)}.$$

ν_2	F \ ν_1	1	2	3	4	5	6	7	8	9
1	0.90	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	0.95	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
	0.975	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3
	0.99	4052	4999	5403	5625	5764	5859	5928	5982	6022
	0.995	16211	20000	21615	22500	23056	23437	23715	23925	24091
	0.999	4053*	5000*	5404*	5625*	5764*	5859*	5929*	5981*	6023*
2	0.90	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	0.95	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
	0.975	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
	0.99	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	0.995	198.5	199.0	199.2	199.2	199.3	199.3	199.4	199.4	199.4
	0.999	998.5	999.0	999.2	999.2	999.3	999.3	999.4	999.4	999.4
3	0.90	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	0.95	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
	0.975	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
	0.99	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
	0.995	55.55	49.80	47.47	46.19	45.39	44.84	44.43	44.13	43.88
	0.999	167.0	148.5	141.1	137.1	134.6	132.8	131.6	130.6	129.9

ν_2	F \ ν_1	10	12	15	20	30	40	60	120	∞
1	0.90	60.19	60.71	61.22	61.74	62.26	62.53	62.79	63.06	63.33
	0.95	241.9	243.9	245.9	248.0	250.1	251.1	252.2	253.3	254.3
	0.975	968.6	976.7	984.9	993.1	1001	1006	1010	1014	1018
	0.99	6056	6106	6157	6209	6261	6287	6313	6339	6366
	0.995	24224	24426	24630	24836	25044	25148	25253	25359	25465
	0.999	6056*	6107*	6158*	6209*	6261*	6287*	6313*	6340*	6366*

带 * 号的数字需乘以 100.

续表

ν_2	F \ ν_1	10	12	15	20	30	40	60	120	∞
2	0.90	9.39	9.41	9.42	9.44	9.46	9.47	9.47	9.48	9.49
	0.95	19.40	19.41	19.43	19.45	19.46	19.47	19.48	19.49	19.50
	0.975	39.40	39.41	39.43	39.45	39.46	39.47	39.48	39.49	39.50
	0.99	99.40	99.42	99.43	99.45	99.47	99.47	99.48	99.49	99.50
	0.995	199.4	199.4	199.4	199.4	199.5	199.5	199.5	199.5	199.5
	0.999	999.4	999.4	999.4	999.4	999.5	999.5	999.5	999.5	999.5
3	0.90	5.23	5.22	5.20	5.18	5.17	5.16	5.15	5.14	5.13
	0.95	8.79	8.74	8.70	8.66	8.62	8.59	8.57	8.55	8.53
	0.975	14.42	14.34	14.25	14.17	14.08	14.04	13.99	13.95	13.90
	0.99	27.23	27.05	26.87	26.69	26.50	26.41	26.32	26.22	26.13
	0.995	43.69	43.39	43.08	42.78	42.47	42.31	42.15	41.99	41.83
	0.999	129.2	128.3	127.4	126.4	125.4	125.0	124.5	124.0	123.5

ν_2	F \ ν_1	1	2	3	4	5	6	7	8	9
4	0.90	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	0.95	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	0.975	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
	0.99	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	0.995	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21.35	21.14
	0.999	74.14	61.25	56.18	53.44	51.71	50.53	49.66	49.00	48.47
5	0.90	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	0.95	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	0.975	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68
	0.99	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	0.995	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77
	0.999	47.18	37.12	33.20	31.09	29.75	28.84	28.16	27.64	27.24
6	0.90	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	0.95	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	0.975	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52
	0.99	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	0.995	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39
	0.999	35.51	27.00	23.70	21.92	20.81	20.03	19.46	19.03	18.69
7	0.90	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	0.95	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	0.975	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82
	0.99	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	0.995	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51
	0.999	29.25	21.69	18.77	17.19	16.21	15.52	15.02	14.63	14.33

续表

ν_2	F \ ν_1	1	2	3	4	5	6	7	8	9
8	0.90	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	0.95	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	0.975	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36
	0.99	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	0.995	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34
	0.999	25.42	18.49	15.83	14.39	13.49	12.86	12.40	12.04	11.77
9	0.90	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	0.95	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	0.975	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03
	0.99	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	0.995	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54
	0.999	22.86	16.39	13.90	12.56	11.71	11.13	10.70	10.37	10.11
10	0.90	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	0.95	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	0.975	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78
	0.99	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	0.995	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97
	0.999	21.04	14.91	12.55	11.28	10.48	9.92	9.52	9.20	8.96
11	0.90	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	0.95	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	0.975	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59
	0.99	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	0.995	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54
	0.999	19.69	13.81	11.56	10.35	9.58	9.05	8.66	8.35	8.12

ν_2	F \ ν_1	10	12	15	20	30	40	60	120	∞
4	0.90	3.92	3.90	3.87	3.84	3.82	3.80	3.79	3.78	3.76
	0.95	5.96	5.91	5.86	5.80	5.75	5.72	5.69	5.66	5.63
	0.975	8.84	8.75	8.66	8.56	8.46	8.41	8.36	8.31	8.26
	0.99	14.55	14.37	14.20	14.02	13.84	13.75	13.65	13.56	13.46
	0.995	20.97	20.70	20.44	20.17	19.89	19.75	19.61	19.47	19.32
	0.999	48.05	47.41	46.76	46.10	45.43	45.09	44.75	44.40	44.05
5	0.90	3.30	3.27	3.24	3.21	3.17	3.16	3.14	3.12	3.10
	0.95	4.74	4.68	4.62	4.56	4.50	4.46	4.43	4.40	4.36
	0.975	6.62	6.52	6.43	6.33	6.23	6.18	6.12	6.07	6.02
	0.99	10.05	9.89	9.72	9.55	9.38	9.29	9.20	9.11	9.02
	0.995	13.62	13.38	13.15	12.90	12.66	12.53	12.40	12.27	12.14
	0.999	26.92	26.42	25.91	25.39	24.87	24.60	24.33	24.06	23.79

续表

ν_2	F \ ν_1	10	12	15	20	30	40	60	120	∞
6	0.90	2.94	2.90	2.87	2.84	2.80	2.78	2.76	2.74	2.72
	0.95	4.06	4.00	3.94	3.87	3.81	3.77	3.74	3.70	3.67
	0.975	5.40	5.37	5.27	5.17	5.07	5.01	4.96	4.90	4.85
	0.99	7.87	7.72	7.56	7.40	7.23	7.14	7.06	6.97	6.88
	0.995	10.25	10.03	9.81	9.59	9.36	9.24	9.12	9.00	8.88
	0.999	18.41	17.99	17.56	17.12	16.67	16.44	16.21	15.99	15.75
7	0.90	2.70	2.67	2.63	2.59	2.56	2.54	2.51	2.49	2.47
	0.95	3.64	3.57	3.51	3.44	3.38	3.34	3.30	3.27	3.23
	0.975	4.76	4.67	4.57	4.47	4.36	4.31	4.25	4.20	4.14
	0.99	6.62	6.47	6.31	6.16	5.99	5.91	5.82	5.74	5.65
	0.995	8.38	8.18	7.97	7.75	7.53	7.42	7.31	7.19	7.08
	0.999	14.08	13.71	13.32	12.93	12.53	12.33	12.12	11.91	11.70
8	0.90	2.54	2.50	2.46	2.42	2.38	2.36	2.34	2.32	2.29
	0.95	3.35	3.28	3.22	3.15	3.08	3.04	3.01	2.97	2.93
	0.975	4.30	4.20	4.10	4.00	3.89	3.84	3.78	3.73	3.67
	0.99	5.81	5.67	5.52	5.36	5.20	5.12	5.03	4.95	4.86
	0.995	7.21	7.01	6.81	6.61	6.40	6.29	6.18	6.06	5.95
	0.999	11.54	11.19	10.84	10.48	10.11	9.92	9.73	9.53	9.33
9	0.90	2.42	2.38	2.34	2.30	2.25	2.23	2.21	2.18	2.16
	0.95	3.14	3.07	3.01	2.94	2.86	2.83	2.79	2.75	2.71
	0.975	3.96	3.87	3.77	3.67	3.56	3.51	3.45	3.39	3.33
	0.99	5.26	5.11	4.96	4.81	4.65	4.57	4.48	4.40	4.31
	0.995	6.42	6.23	6.03	5.83	5.62	5.52	5.41	5.30	5.19
	0.999	9.89	9.57	9.24	8.90	8.55	8.37	8.19	8.00	7.81
10	0.90	2.32	2.28	2.24	2.20	2.16	2.13	2.11	2.08	2.06
	0.95	2.98	2.91	2.85	2.77	2.70	2.66	2.62	2.58	2.54
	0.975	3.72	3.62	3.52	3.42	3.31	3.26	3.20	3.14	3.08
	0.99	4.85	4.71	4.56	4.41	4.25	4.17	4.08	4.00	3.91
	0.995	5.85	5.66	5.47	5.27	5.07	4.97	4.86	4.75	4.64
	0.999	8.75	8.45	8.13	7.80	7.47	7.30	7.12	6.94	6.76
11	0.90	2.25	2.21	2.17	2.12	2.08	2.05	2.03	2.00	1.97
	0.95	2.85	2.79	2.72	2.65	2.57	2.53	2.49	2.45	2.40
	0.975	3.53	3.43	3.33	3.23	3.12	3.06	3.00	2.94	2.88
	0.99	4.54	4.40	4.25	4.10	3.94	3.86	3.78	3.69	3.60
	0.995	5.42	5.24	5.05	4.86	4.65	4.55	4.44	4.34	4.23
	0.999	7.92	7.63	7.32	7.01	6.68	6.52	6.35	6.17	6.00

续表

ν_2	F \ ν_1	1	2	3	4	5	6	7	8	9
12	0.90	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	0.95	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	0.975	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44
	0.99	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	0.995	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20
	0.999	18.64	12.97	10.80	9.63	8.89	8.38	8.00	7.71	7.48
13	0.90	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	0.95	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	0.975	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31
	0.99	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	0.995	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94
	0.999	17.81	12.31	10.21	9.07	8.35	7.86	7.49	7.21	6.98
14	0.90	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	0.95	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	0.975	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21
	0.99	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	0.995	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72
	0.999	17.14	11.78	9.73	8.62	7.92	7.43	7.08	6.80	6.58
15	0.90	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	0.95	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	0.975	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
	0.99	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	0.995	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54
	0.999	16.59	11.34	9.34	8.25	7.57	7.09	6.74	6.47	6.26
16	0.90	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	0.95	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	0.975	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05
	0.99	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	0.995	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38
	0.999	16.12	10.97	9.00	7.94	7.27	6.81	6.46	6.19	5.98
17	0.90	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	0.95	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	0.975	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98
	0.99	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
	0.995	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25
	0.999	15.72	10.66	8.73	7.68	7.02	7.56	6.22	5.96	5.75

续表

ν_2	F \ ν_1	1	2	3	4	5	6	7	8	9
18	0.90	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
	0.95	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	0.975	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93
	0.99	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	0.995	10.22	7.21	6.03	5.37	4.96	4.66	4.44	4.28	4.14
	0.999	15.38	10.39	8.49	7.46	6.81	6.35	6.02	5.76	5.56
19	0.90	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
	0.95	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
	0.975	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88
	0.99	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	0.995	10.07	7.09	5.92	5.27	4.85	4.56	4.34	4.18	4.04
	0.999	15.08	10.16	8.28	7.26	6.62	6.18	5.85	5.59	5.39

ν_2	F \ ν_1	10	12	15	20	30	40	60	120	∞
12	0.90	2.19	2.15	2.10	2.06	2.01	1.99	1.96	1.93	1.90
	0.95	2.75	2.69	2.62	2.54	2.47	2.43	2.38	2.34	2.30
	0.975	3.37	3.28	3.18	3.07	2.96	2.91	2.85	2.79	2.72
	0.99	4.30	4.16	4.01	3.86	3.70	3.62	3.54	3.45	3.36
	0.995	5.09	4.91	4.72	4.53	4.33	4.23	4.12	4.01	3.90
	0.999	7.29	7.00	6.71	6.40	6.09	5.93	5.76	5.59	5.42
13	0.90	2.14	2.10	2.05	2.01	1.96	1.93	1.90	1.88	1.85
	0.95	2.67	2.60	2.53	2.46	2.38	2.34	2.30	2.25	2.21
	0.975	3.25	3.15	3.05	2.95	2.84	2.78	2.72	2.66	2.60
	0.99	4.10	3.96	3.82	3.66	3.51	3.43	3.34	3.25	3.17
	0.995	4.82	4.64	4.46	4.27	4.07	3.97	3.87	3.76	3.65
	0.999	6.80	6.52	6.23	5.93	5.63	5.47	5.30	5.14	4.97
14	0.90	2.10	2.05	2.01	1.96	1.91	1.89	1.86	1.83	1.80
	0.95	2.60	2.53	2.46	2.39	2.31	2.27	2.22	2.18	2.13
	0.975	3.15	3.05	2.95	2.84	2.73	2.67	2.61	2.55	2.49
	0.99	3.94	3.80	3.66	3.51	3.35	3.27	3.18	3.09	3.00
	0.995	4.60	4.43	4.25	4.06	3.86	3.76	3.66	3.55	3.44
	0.999	6.40	6.13	5.85	5.56	5.25	5.10	4.94	4.77	4.60
15	0.90	2.06	2.02	1.97	1.92	1.87	1.85	1.82	1.79	1.76
	0.95	2.54	2.48	2.40	2.33	2.25	2.20	2.16	2.11	2.07
	0.975	3.06	2.96	2.86	2.76	2.64	2.59	2.52	2.46	2.40
	0.99	3.80	3.67	3.52	3.37	3.21	3.13	3.05	2.96	2.87
	0.995	4.42	4.25	4.07	3.88	3.69	3.58	3.48	3.37	3.26

续表

ν_2	F \ ν_1	10	12	15	20	30	40	60	120	∞
	0.999	6.08	5.81	5.54	5.25	4.95	4.80	4.64	4.47	4.31
16	0.90	2.03	1.99	1.94	1.89	1.84	1.81	1.78	1.75	1.72
	0.95	2.49	2.42	2.35	2.28	2.19	2.15	2.11	2.06	2.01
	0.975	2.99	2.89	2.79	2.68	2.57	2.51	2.45	2.38	2.32
	0.99	3.69	3.55	3.41	3.26	3.10	3.02	2.93	2.84	2.75
	0.995	4.27	4.10	3.92	3.73	3.54	3.44	3.33	3.22	3.11
	0.999	5.81	5.55	5.27	4.99	4.70	4.54	4.39	4.23	4.06
17	0.90	2.00	1.96	1.91	1.86	1.81	1.78	1.75	1.72	1.69
	0.95	2.45	2.38	2.31	2.23	2.15	2.10	2.06	2.01	1.96
	0.975	2.92	2.82	2.72	2.62	2.50	2.44	2.38	2.32	2.25
	0.99	3.59	3.46	3.31	3.16	3.00	2.92	2.83	2.75	2.65
	0.995	4.14	3.97	3.79	3.61	3.41	3.31	3.21	3.10	2.95
	0.999	5.58	5.32	5.05	4.78	4.48	4.33	4.18	4.02	3.85
18	0.90	1.98	1.93	1.89	1.84	1.78	1.75	1.72	1.69	1.66
	0.95	2.41	2.34	2.27	2.19	2.11	2.06	2.02	1.97	1.92
	0.975	2.87	2.77	2.67	2.56	2.44	2.38	2.32	2.26	2.19
	0.99	3.51	3.37	3.23	3.08	2.92	2.84	2.75	2.66	2.57
	0.995	4.03	3.86	3.68	3.50	3.30	3.20	3.10	2.99	2.87
	0.999	5.39	5.13	4.87	4.59	4.30	4.15	4.00	3.84	3.67
19	0.90	1.96	1.91	1.86	1.81	1.76	1.73	1.70	1.67	1.63
	0.95	2.38	2.31	2.23	2.16	2.07	2.03	1.98	1.93	1.88
	0.975	2.82	2.72	2.62	2.51	2.39	2.33	2.27	2.20	2.13
	0.99	3.43	3.30	3.15	3.00	2.84	2.76	2.67	2.58	2.49
	0.995	3.93	3.76	3.59	3.40	3.21	3.11	3.00	2.89	2.78
	0.999	5.22	4.97	4.70	4.43	4.14	3.99	3.84	3.68	3.51

ν_2	F \ ν_1	1	2	3	4	5	6	7	8	9
20	0.90	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	0.95	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	0.975	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84
	0.99	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	0.995	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96
	0.999	14.82	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24
21	0.90	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
	0.95	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	0.975	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80
	0.99	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	0.995	9.83	6.89	5.73	5.09	4.68	4.39	4.18	4.01	3.88

续表

ν_2 \ F \ ν_1		1	2	3	4	5	6	7	8	9
	0.999	14.59	9.77	7.94	6.95	6.32	5.88	5.56	5.31	5.11
22	0.90	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	0.95	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	0.975	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76
	0.99	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	0.995	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81
	0.999	14.38	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99
23	0.90	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	0.95	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	0.975	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73
	0.99	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	0.995	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.75
	0.999	14.19	9.47	7.67	6.69	6.08	5.65	5.33	5.09	4.89
24	0.90	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	0.95	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	0.975	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70
	0.99	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	0.995	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69
	0.999	14.03	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.80
25	0.90	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	0.95	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	0.975	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68
	0.99	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	0.995	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64
	0.999	13.88	9.22	7.45	6.49	5.88	5.46	5.15	4.91	4.71
26	0.90	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	0.95	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	0.975	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65
	0.99	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	0.995	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.60
	0.999	13.74	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.64
27	0.90	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	0.95	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	0.975	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63
	0.99	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	0.995	9.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.56

续表

ν_2	F \ ν_1	1	2	3	4	5	6	7	8	9
	0.999	13.61	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4.57

ν_2	F \ ν_1	10	12	15	20	30	40	60	120	∞
20	0.90	1.94	1.89	1.84	1.79	1.74	1.71	1.68	1.64	1.61
	0.95	2.35	2.28	2.20	2.12	2.04	1.99	1.95	1.90	1.84
	0.975	2.77	2.68	2.57	2.46	2.35	2.29	2.22	2.16	2.09
	0.99	3.37	3.23	3.09	2.94	2.78	2.69	2.61	2.52	2.42
	0.995	3.85	3.68	3.50	3.32	3.12	3.02	2.92	2.81	2.69
	0.999	5.08	4.82	4.56	4.29	4.00	3.86	3.70	3.54	3.38
21	0.90	1.92	1.87	1.83	1.78	1.72	1.69	1.66	1.62	1.59
	0.95	2.32	2.25	2.18	2.10	2.01	1.96	1.92	1.87	1.81
	0.975	2.73	2.64	2.53	2.42	2.31	2.25	2.18	2.11	2.04
	0.99	3.31	3.17	3.03	2.88	2.72	2.64	2.55	2.46	2.36
	0.995	3.77	3.60	3.43	3.24	3.05	2.95	2.84	2.73	2.61
	0.999	4.95	4.70	4.44	4.17	3.88	3.74	3.58	3.42	3.26
22	0.90	1.90	1.86	1.81	1.76	1.70	1.67	1.64	1.60	1.57
	0.95	2.30	2.23	2.15	2.07	1.98	1.94	1.89	1.84	1.78
	0.975	2.70	2.60	2.50	2.39	2.27	2.21	2.14	2.08	2.00
	0.99	3.26	3.12	2.98	2.83	2.67	2.58	2.50	2.40	2.31
	0.995	3.70	3.54	3.36	3.18	2.98	2.88	2.77	2.66	2.55
	0.999	4.83	4.58	4.33	4.06	3.78	3.63	3.48	3.32	3.15
23	0.90	1.89	1.84	1.80	1.74	1.69	1.66	1.62	1.59	1.55
	0.95	2.27	2.20	2.13	2.05	1.96	1.91	1.86	1.81	1.76
	0.975	2.67	2.57	2.47	2.36	2.24	2.18	2.11	2.04	1.97
	0.99	3.21	3.07	2.93	2.78	2.62	2.54	2.45	2.35	2.26
	0.995	3.64	3.47	3.30	3.12	2.92	2.82	2.71	2.60	2.48
	0.999	4.73	4.48	4.23	3.96	3.68	3.53	3.38	3.22	3.05
24	0.90	1.88	1.83	1.78	1.73	1.67	1.64	1.61	1.57	1.53
	0.95	2.25	2.18	2.11	2.03	1.94	1.89	1.84	1.79	1.73
	0.975	2.64	2.54	2.44	2.33	2.21	2.15	2.08	2.01	1.94
	0.99	3.17	3.03	2.89	2.74	2.58	2.49	2.40	2.31	2.21
	0.995	3.59	3.42	3.25	3.06	2.87	2.77	2.66	2.55	2.43
	0.999	4.64	4.39	4.14	3.87	3.59	3.45	3.29	3.14	2.97
25	0.90	1.87	1.82	1.77	1.72	1.66	1.63	1.59	1.56	1.52
	0.95	2.24	2.16	2.09	2.01	1.92	1.87	1.82	1.77	1.71
	0.975	2.61	2.51	2.41	2.30	2.18	2.12	2.05	1.98	1.91

续表

ν_2	F \ ν_1	10	12	15	20	30	40	60	120	∞
	0.99	3.13	2.99	2.85	2.70	2.54	2.45	2.36	2.27	2.17
	0.995	3.54	3.37	3.20	3.01	2.82	2.72	2.61	2.50	2.38
	0.999	4.56	4.31	4.06	3.79	3.52	3.37	3.22	3.06	2.89
26	0.90	1.86	1.81	1.76	1.71	1.65	1.61	1.58	1.54	1.50
	0.95	2.22	2.15	2.07	1.99	1.90	1.85	1.80	1.75	1.69
	0.975	2.59	2.49	2.39	2.28	2.16	2.09	2.03	1.95	1.88
	0.99	3.09	2.96	2.81	2.66	2.50	2.42	2.33	2.23	2.13
	0.995	3.49	3.33	3.15	2.97	2.77	2.67	2.56	2.45	2.33
	0.999	4.48	4.24	3.99	3.72	3.44	3.30	3.15	2.99	2.82
27	0.90	1.85	1.80	1.75	1.70	1.64	1.60	1.57	1.53	1.49
	0.95	2.20	2.13	2.06	1.97	1.88	1.84	1.79	1.73	1.67
	0.975	2.57	2.47	2.36	2.25	2.13	2.07	2.00	1.93	1.85
	0.99	3.06	2.93	2.78	2.63	2.47	2.38	2.29	2.20	2.10
	0.995	3.45	3.28	3.11	2.93	2.73	2.63	2.52	2.41	2.29
	0.999	4.41	4.17	3.92	3.66	3.38	3.23	3.08	2.92	2.75

ν_2	F \ ν_1	1	2	3	4	5	6	7	8	9
28	0.90	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	0.95	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	0.975	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61
	0.99	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	0.995	9.28	6.44	5.32	4.70	4.30	4.02	3.81	3.65	3.52
	0.999	13.50	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.50
29	0.90	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	0.95	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	0.975	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59
	0.99	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	0.995	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.48
	0.999	13.39	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.45
30	0.90	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	0.95	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	0.975	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	0.99	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	0.995	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45
	0.999	13.29	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39
40	0.90	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79

续表

ν_2	F \ ν_1	1	2	3	4	5	6	7	8	9
	0.95	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	0.975	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45
	0.99	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	0.995	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22
	0.999	12.61	8.25	6.60	5.70	5.13	4.73	4.44	4.21	4.02
60	0.90	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	0.95	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	0.975	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	0.99	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	0.995	8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01
	0.999	11.97	7.76	6.17	5.31	4.76	4.37	4.09	3.87	3.69
120	0.90	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68
	0.95	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
	0.975	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22
	0.99	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
	0.995	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81
	0.999	11.38	7.32	5.79	4.95	4.42	4.04	3.77	3.55	3.38
∞	0.90	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63
	0.95	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88
	0.975	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11
	0.99	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41
	0.995	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62
	0.999	10.83	6.91	5.42	4.62	4.10	3.74	3.47	3.27	3.10

ν_2	F \ ν_1	10	12	15	20	30	40	60	120	∞
28	0.90	1.84	1.79	1.74	1.69	1.63	1.59	1.56	1.52	1.48
	0.95	2.19	2.12	2.04	1.96	1.87	1.82	1.77	1.71	1.65
	0.975	2.55	2.45	2.34	2.23	2.11	2.05	1.98	1.91	1.83
	0.99	3.03	2.90	2.75	2.60	2.44	2.35	2.26	2.17	2.06
	0.995	3.41	3.25	3.07	2.89	2.69	2.59	2.48	2.37	2.25
	0.999	4.35	4.11	3.86	3.60	3.32	3.18	3.02	2.86	2.69
29	0.90	1.83	1.78	1.73	1.68	1.62	1.58	1.55	1.51	1.47
	0.95	2.18	2.10	2.03	1.94	1.85	1.81	1.75	1.70	1.64
	0.975	2.53	2.43	2.32	2.21	2.09	2.03	1.96	1.89	1.81
	0.99	3.00	2.87	2.73	2.57	2.41	2.33	2.23	2.14	2.03

续表

ν_2	F \ ν_1	10	12	15	20	30	40	60	120	∞
	0.995	3.38	3.21	3.04	2.86	2.66	2.56	2.45	2.33	2.21
	0.999	4.29	4.05	3.80	3.54	3.27	3.12	2.97	2.81	2.64
30	0.90	1.82	1.77	1.72	1.67	1.61	1.57	1.54	1.50	1.46
	0.95	2.16	2.09	2.01	1.93	1.84	1.79	1.74	1.68	1.62
	0.975	2.51	2.41	2.31	2.20	2.07	2.01	1.94	1.87	1.79
	0.99	2.98	2.84	2.70	2.55	2.39	2.30	2.21	2.11	2.01
	0.995	3.34	3.18	3.01	2.82	2.63	2.52	2.42	2.30	2.18
	0.999	4.24	4.00	3.75	3.49	3.22	3.07	2.92	2.76	2.59
40	0.90	1.76	1.71	1.66	1.61	1.54	1.51	1.47	1.42	1.38
	0.95	2.08	2.00	1.92	1.84	1.74	1.69	1.64	1.58	1.51
	0.975	2.39	2.29	2.18	2.07	1.94	1.88	1.80	1.72	1.64
	0.99	2.80	2.66	2.52	2.37	2.20	2.11	2.02	1.92	1.80
	0.995	3.12	2.95	2.78	2.60	2.40	2.30	2.18	2.06	1.93
	0.999	3.87	3.64	3.40	3.15	2.87	2.73	2.57	2.41	2.23
60	0.90	1.71	1.66	1.60	1.54	1.48	1.44	1.40	1.35	1.29
	0.95	1.99	1.92	1.84	1.75	1.65	1.59	1.53	1.47	1.39
	0.975	2.27	2.17	2.06	1.94	1.82	1.74	1.67	1.58	1.48
	0.99	2.63	2.50	2.35	2.20	2.03	1.94	1.84	1.73	1.60
	0.995	2.90	2.74	2.57	2.39	2.19	2.08	1.96	1.83	1.69
	0.999	3.54	3.31	3.08	2.83	2.55	2.41	2.25	2.08	1.89
120	0.90	1.65	1.60	1.55	1.48	1.41	1.37	1.32	1.26	1.19
	0.95	1.91	1.83	1.75	1.66	1.55	1.50	1.43	1.35	1.25
	0.975	2.16	2.05	1.94	1.82	1.69	1.61	1.53	1.43	1.31
	0.99	2.47	2.34	2.19	2.03	1.86	1.76	1.66	1.53	1.38
	0.995	2.71	2.54	2.37	2.19	1.98	1.87	1.75	1.61	1.43
	0.999	3.24	3.02	2.78	2.53	2.26	2.11	1.95	1.76	1.54
∞	0.90	1.60	1.55	1.49	1.42	1.34	1.30	1.24	1.17	1.00
	0.95	1.83	1.75	1.67	1.57	1.46	1.39	1.32	1.22	1.00
	0.975	2.05	1.94	1.83	1.71	1.57	1.48	1.39	1.27	1.00
	0.99	2.32	2.18	2.04	1.88	1.70	1.59	1.47	1.32	1.00
	0.995	2.52	2.36	2.19	2.00	1.79	1.67	1.53	1.36	1.00
	0.999	2.96	2.74	2.51	2.27	1.99	1.84	1.66	1.45	1.00

表 10.1 置信水平 68.27%

n_0 \ b	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0
0	0.00 1.29	0.00 0.80	0.00 0.54	0.00 0.41	0.00 0.41	0.00 0.25	0.00 0.25	0.00 0.21	0.00 0.21	0.00 0.19
1	0.37 2.75	0.00 2.25	0.00 1.75	0.00 1.32	0.00 0.97	0.00 0.68	0.00 0.50	0.00 0.50	0.00 0.36	0.00 0.30
2	0.74 4.25	0.44 3.75	0.14 3.25	0.00 2.75	0.00 2.25	0.00 1.80	0.00 1.41	0.00 1.09	0.00 0.81	0.00 0.47
3	1.10 5.30	0.80 4.80	0.54 4.30	0.32 3.80	0.00 3.30	0.00 2.80	0.00 2.30	0.00 1.84	0.00 1.45	0.00 0.91
4	2.34 6.78	1.84 6.28	1.34 5.78	0.91 5.28	0.44 4.78	0.25 4.28	0.00 3.78	0.00 3.28	0.00 2.78	0.00 1.90
5	2.75 7.81	2.25 7.31	1.75 6.81	1.32 6.31	0.97 5.81	0.68 5.31	0.45 4.81	0.20 4.31	0.00 3.81	0.00 2.81
6	3.82 9.28	3.32 8.78	2.82 8.28	2.32 7.78	1.82 7.28	1.37 6.78	1.01 6.28	0.62 5.78	0.36 5.28	0.00 4.28
7	4.25 10.30	3.75 9.80	3.25 9.30	2.75 8.80	2.25 8.30	1.80 7.80	1.41 7.30	1.09 6.80	0.81 6.30	0.32 5.30
8	5.30 11.32	4.80 10.82	4.30 10.32	3.80 9.82	3.30 9.32	2.80 8.82	2.30 8.32	1.84 7.82	1.45 7.32	0.82 6.32
9	6.33 12.79	5.83 12.29	5.33 11.79	4.83 11.29	4.33 10.79	3.83 10.29	3.33 9.79	2.83 9.29	2.33 8.79	1.44 7.79
10	6.78 13.81	6.28 13.31	5.78 12.81	5.28 12.31	4.78 11.81	4.28 11.31	3.78 10.81	3.28 10.31	2.78 9.81	1.90 8.81
11	7.81 14.82	7.31 14.32	6.81 13.82	6.31 13.32	5.81 12.82	5.31 12.32	4.81 11.82	4.31 11.32	3.81 10.82	2.81 9.82
12	8.83 16.29	8.33 15.79	7.83 15.29	7.33 14.79	6.83 14.29	6.33 13.79	5.83 13.29	5.33 12.79	4.83 12.29	3.83 11.29
13	9.28 17.30	8.78 16.80	8.28 16.30	7.78 15.80	7.28 15.30	6.78 14.80	6.28 14.30	5.78 13.80	5.28 13.30	4.28 12.30
14	10.30 18.32	9.80 17.82	9.30 17.32	8.80 16.82	8.30 16.32	7.80 15.82	7.30 15.32	6.80 14.82	6.30 14.32	5.30 13.32
15	11.32 19.32	10.82 18.82	10.32 18.32	9.82 17.82	9.32 17.32	8.82 16.82	8.32 16.32	7.82 15.82	7.32 15.32	6.32 14.32
16	12.33 20.80	11.83 20.30	11.33 19.80	10.83 19.30	10.33 18.80	9.83 18.30	9.33 17.80	8.83 17.30	8.33 16.80	7.33 15.80
17	12.79 21.81	12.29 21.31	11.79 20.81	11.29 20.31	10.79 19.81	10.29 19.31	9.79 18.81	9.29 18.31	8.79 17.81	7.79 16.81
18	13.81 22.82	13.31 22.32	12.81 21.82	12.31 21.32	11.81 20.82	11.31 20.32	10.81 19.82	10.31 19.32	9.81 18.82	8.81 17.82
19	14.82 23.82	14.32 23.32	13.82 22.82	13.32 22.32	12.82 21.82	12.32 21.32	11.82 20.82	11.32 20.32	10.82 19.82	9.82 18.82
20	15.83 25.30	15.33 24.80	14.83 24.30	14.33 23.80	13.83 23.30	13.33 22.80	12.83 22.30	12.33 21.80	11.83 21.30	10.83 20.30

表 10.1～ 表 10.4 似然比顺序求和方法求得的信号区间内信号泊松事例数期望值 μ 的置信区间 $[\mu_1, \mu_2]$. 其中, n_0 和 b 是信号区间内观测到的事例总数和已知平均本底事例数 (0～15).

续表

n_0 \ b	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0
0	0.00 0.18	0.00 0.17	0.00 0.17	0.00 0.17	0.00 0.16	0.00 0.16	0.00 0.16	0.00 0.16	0.00 0.16	0.00 0.15
1	0.00 0.24	0.00 0.21	0.00 0.20	0.00 0.19	0.00 0.18	0.00 0.17	0.00 0.17	0.00 0.17	0.00 0.17	0.00 0.16
2	0.00 0.31	0.00 0.27	0.00 0.23	0.00 0.21	0.00 0.20	0.00 0.19	0.00 0.19	0.00 0.18	0.00 0.18	0.00 0.18
3	0.00 0.69	0.00 0.42	0.00 0.31	0.00 0.26	0.00 0.23	0.00 0.22	0.00 0.21	0.00 0.20	0.00 0.20	0.00 0.19
4	0.00 1.22	0.00 0.69	0.00 0.60	0.00 0.38	0.00 0.30	0.00 0.26	0.00 0.24	0.00 0.23	0.00 0.22	0.00 0.21
5	0.00 1.92	0.00 1.23	0.00 0.99	0.00 0.60	0.00 0.48	0.00 0.35	0.00 0.29	0.00 0.26	0.00 0.24	0.00 0.23
6	0.00 3.28	0.00 2.38	0.00 1.65	0.00 1.06	0.00 0.63	0.00 0.53	0.00 0.42	0.00 0.33	0.00 0.29	0.00 0.26
7	0.00 4.30	0.00 3.30	0.00 2.40	0.00 1.66	0.00 1.07	0.00 0.88	0.00 0.53	0.00 0.47	0.00 0.38	0.00 0.32
8	0.31 5.32	0.00 4.32	0.00 3.32	0.00 2.41	0.00 1.67	0.00 1.46	0.00 0.94	0.00 0.62	0.00 0.48	0.00 0.43
9	0.69 6.79	0.27 5.79	0.00 4.79	0.00 3.79	0.00 2.87	0.00 2.10	0.00 1.46	0.00 0.94	0.00 0.78	0.00 0.50
10	1.22 7.81	0.69 6.81	0.23 5.81	0.00 4.81	0.00 3.81	0.00 2.89	0.00 2.11	0.00 1.47	0.00 1.03	0.00 0.84
11	1.92 8.82	1.23 7.82	0.60 6.82	0.19 5.82	0.00 4.82	0.00 3.82	0.00 2.90	0.00 2.12	0.00 1.54	0.00 1.31
12	2.83 10.29	1.94 9.29	1.12 8.29	0.60 7.29	0.12 6.29	0.00 5.29	0.00 4.29	0.00 3.36	0.00 2.57	0.00 1.89
13	3.28 11.30	2.38 10.30	1.65 9.30	1.06 8.30	0.60 7.30	0.05 6.30	0.00 5.30	0.00 4.30	0.00 3.37	0.00 2.57
14	4.30 12.32	3.30 11.32	2.40 10.32	1.66 9.32	1.07 8.32	0.53 7.32	0.00 6.32	0.00 5.32	0.00 4.32	0.00 3.38
15	5.32 13.32	4.32 12.32	3.32 11.32	2.41 10.32	1.67 9.32	1.00 8.32	0.53 7.32	0.00 6.32	0.00 5.32	0.00 4.32
16	6.33 14.80	5.33 13.80	4.33 12.80	3.33 11.80	2.43 10.80	1.46 9.80	0.94 8.80	0.47 7.80	0.00 6.80	0.00 5.80
17	6.79 15.81	5.79 14.81	4.79 13.81	3.79 12.81	2.87 11.81	2.10 10.81	1.46 9.81	0.94 8.81	0.48 7.81	0.00 6.81
18	7.81 16.82	6.81 15.82	5.81 14.82	4.81 13.82	3.81 12.82	2.89 11.82	2.11 10.82	1.47 9.82	0.93 8.82	0.43 7.82
19	8.82 17.82	7.82 16.82	6.82 15.82	5.82 14.82	4.82 13.82	3.82 12.82	2.90 11.82	2.12 10.82	1.48 9.82	0.84 8.82
20	9.83 19.30	8.83 18.30	7.83 17.30	6.83 16.30	5.83 15.30	4.83 14.30	3.83 13.30	2.91 12.30	2.12 11.30	1.31 10.30

表 10.2　置信水平 90%

n_0 \ b	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0
0	0.00 2.44	0.00 1.94	0.00 1.61	0.00 1.33	0.00 1.26	0.00 1.18	0.00 1.08	0.00 1.06	0.00 1.01	0.00 0.98
1	0.11 4.36	0.00 3.86	0.00 3.36	0.00 2.91	0.00 2.53	0.00 2.19	0.00 1.88	0.00 1.59	0.00 1.39	0.00 1.22
2	0.53 5.91	0.03 5.41	0.00 4.91	0.00 4.41	0.00 3.91	0.00 3.45	0.00 3.04	0.00 2.67	0.00 2.33	0.00 1.73
3	1.10 7.42	0.60 6.92	0.10 6.42	0.00 5.92	0.00 5.42	0.00 4.92	0.00 4.42	0.00 3.95	0.00 3.53	0.00 2.78
4	1.47 8.60	1.17 8.10	0.74 7.60	0.24 7.10	0.00 6.60	0.00 6.10	0.00 5.60	0.00 5.10	0.00 4.60	0.00 3.60
5	1.84 9.99	1.53 9.49	1.25 8.99	0.93 8.49	0.43 7.99	0.00 7.49	0.00 6.99	0.00 6.49	0.00 5.99	0.00 4.99
6	2.21 11.47	1.61 10.47	1.61 10.47	1.33 9.97	1.08 9.47	0.65 8.97	0.15 8.47	0.00 7.97	0.00 7.47	0.00 6.47
7	3.56 12.53	2.56 11.53	2.56 11.53	2.09 11.03	1.59 10.53	1.18 10.03	0.89 9.53	0.39 9.03	0.00 8.53	0.00 7.53
8	3.96 13.99	2.96 12.99	2.96 12.99	2.51 12.49	2.14 11.99	1.81 11.49	1.51 10.99	1.06 10.49	0.66 9.99	0.00 8.99
9	4.36 15.30	3.36 14.30	3.36 14.30	2.91 13.80	2.53 13.30	2.19 12.80	1.88 12.30	1.59 11.80	1.33 11.30	0.43 10.30
10	5.50 16.50	4.50 15.50	4.50 15.50	4.00 15.00	3.50 14.50	3.04 14.00	2.63 13.50	2.27 13.00	1.94 12.50	1.19 11.50
11	5.91 17.81	4.91 16.81	4.91 16.81	4.41 16.31	3.91 15.81	3.45 15.31	3.04 14.81	2.67 14.31	2.33 13.81	1.73 12.81
12	7.01 19.00	6.01 18.00	6.01 18.00	5.51 17.50	5.01 17.00	4.51 16.50	4.01 16.00	3.54 15.50	3.12 15.00	2.38 14.00
13	7.42 20.05	6.42 19.05	6.42 19.05	5.92 18.55	5.42 18.05	4.92 17.55	4.42 17.05	3.95 16.55	3.53 16.05	2.78 15.05
14	8.50 21.50	7.50 20.50	7.50 20.50	7.00 20.00	6.50 19.50	6.00 19.00	5.50 18.50	5.00 18.00	4.50 17.50	3.59 16.50
15	9.48 22.52	8.48 21.52	8.48 21.52	7.98 21.02	7.48 20.52	6.98 20.02	6.48 19.52	5.98 19.02	5.48 18.52	4.48 17.52
16	9.99 23.99	8.99 22.99	8.99 22.99	8.49 22.49	7.99 21.99	7.49 21.49	6.99 20.99	6.49 20.49	5.99 19.99	4.99 18.99
17	11.04 25.02	10.54 24.52	9.54 23.52	9.54 23.52	9.04 23.02	8.54 22.52	8.04 22.02	7.54 21.52	7.04 21.02	6.04 20.02
18	11.47 26.16	10.97 25.66	9.97 24.66	9.97 24.66	9.47 24.16	8.97 23.66	8.47 23.16	7.97 22.66	7.47 22.16	6.47 21.16
19	12.51 27.51	12.01 27.01	11.01 26.01	11.01 26.01	10.51 25.51	10.01 25.01	9.51 24.51	9.01 24.01	8.51 23.51	7.51 22.51
20	13.55 28.52	13.05 28.02	12.05 27.02	12.05 27.02	11.55 26.52	11.05 26.02	10.55 25.52	10.05 25.02	9.55 24.52	8.55 23.52

续表

n_0 \ b	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0
0	0.00 0.97	0.00 0.95	0.00 0.94	0.00 0.94	0.00 0.93	0.00 0.93	0.00 0.92	0.00 0.92	0.00 0.92	0.00 0.92
1	0.00 1.14	0.00 1.10	0.00 1.07	0.00 1.03	0.00 1.03	0.00 1.01	0.00 1.00	0.00 0.99	0.00 0.99	0.00 0.98
2	0.00 1.57	0.00 1.38	0.00 1.27	0.00 1.21	0.00 1.15	0.00 1.11	0.00 1.09	0.00 1.08	0.00 1.06	0.00 1.05
3	0.00 2.14	0.00 1.75	0.00 1.49	0.00 1.37	0.00 1.29	0.00 1.24	0.00 1.21	0.00 1.18	0.00 1.15	0.00 1.14
4	0.00 2.83	0.00 2.56	0.00 1.98	0.00 1.57	0.00 1.57	0.00 1.45	0.00 1.37	0.00 1.31	0.00 1.27	0.00 1.24
5	0.00 4.07	0.00 3.28	0.00 2.60	0.00 2.38	0.00 1.85	0.00 1.70	0.00 1.58	0.00 1.48	0.00 1.39	0.00 1.32
6	0.00 5.47	0.00 4.54	0.00 3.73	0.00 3.02	0.00 2.40	0.00 2.21	0.00 1.86	0.00 1.67	0.00 1.55	0.00 1.47
7	0.00 6.53	0.00 5.53	0.00 4.58	0.00 3.77	0.00 3.26	0.00 2.81	0.00 2.23	0.00 2.07	0.00 1.86	0.00 1.69
8	0.00 7.99	0.00 6.99	0.00 5.99	0.00 5.05	0.00 4.22	0.00 3.49	0.00 2.83	0.00 2.62	0.00 2.11	0.00 1.95
9	0.00 9.30	0.00 8.30	0.00 7.30	0.00 6.30	0.00 5.30	0.00 4.30	0.00 3.93	0.00 3.25	0.00 2.64	0.00 2.45
10	0.22 10.50	0.00 9.50	0.00 8.50	0.00 7.50	0.00 6.50	0.00 5.56	0.00 4.71	0.00 3.95	0.00 3.27	0.00 3.00
11	1.01 11.81	0.02 10.81	0.00 9.81	0.00 8.81	0.00 7.81	0.00 6.81	0.00 5.81	0.00 4.81	0.00 4.39	0.00 3.69
12	1.57 13.00	0.83 12.00	0.00 11.00	0.00 10.00	0.00 9.00	0.00 8.00	0.00 7.00	0.00 6.05	0.00 5.19	0.00 4.42
13	2.14 14.05	1.50 13.05	0.65 12.05	0.00 11.05	0.00 10.05	0.00 9.05	0.00 8.05	0.00 7.05	0.00 6.08	0.00 5.22
14	2.83 15.50	2.13 14.50	1.39 13.50	0.47 12.50	0.00 11.50	0.00 10.50	0.00 9.50	0.00 8.50	0.00 7.50	0.00 6.55
15	3.48 16.52	2.56 15.52	1.98 14.52	1.26 13.52	0.30 12.52	0.00 11.52	0.00 10.52	0.00 9.52	0.00 8.52	0.00 7.52
16	4.07 17.99	3.28 16.99	2.60 15.99	1.82 14.99	1.13 13.99	0.14 12.99	0.00 11.99	0.00 10.99	0.00 9.99	0.00 8.99
17	5.04 19.02	4.11 18.02	3.32 17.02	2.38 16.02	1.81 15.02	0.98 14.02	0.00 13.02	0.00 12.02	0.00 11.02	0.00 10.02
18	5.47 20.16	4.54 19.16	3.73 18.16	3.02 17.16	2.40 16.16	1.70 15.16	0.82 14.16	0.00 13.16	0.00 12.16	0.00 11.16
19	6.51 21.51	5.51 20.51	4.58 19.51	3.77 18.51	3.05 17.51	2.21 16.51	1.58 15.51	0.67 14.51	0.00 13.51	0.00 12.51
20	7.55 22.52	6.55 21.52	5.55 20.52	4.55 19.52	3.55 18.52	2.81 17.52	2.23 16.52	1.48 15.52	0.53 14.52	0.00 13.52

表 10.3 置信水平 95%

n_0 \ b	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0
0	0.00 3.09	0.00 2.63	0.00 2.33	0.00 1.78	0.00 1.63	0.00 1.78	0.00 1.63	0.00 1.63	0.00 1.57	0.00 1.54
1	0.05 5.14	0.00 4.64	0.00 4.14	0.00 3.30	0.00 3.30	0.00 2.95	0.00 2.63	0.00 2.33	0.00 2.08	0.00 1.88
2	0.36 6.72	0.00 6.22	0.00 5.72	0.00 5.22	0.00 4.72	0.00 4.25	0.00 3.84	0.00 3.46	0.00 3.11	0.00 2.49
3	0.82 8.25	0.32 7.75	0.00 7.25	0.00 6.75	0.00 6.25	0.00 5.75	0.00 5.25	0.00 4.78	0.00 4.35	0.00 3.58
4	1.37 9.76	0.87 9.26	0.37 8.76	0.00 8.26	0.00 7.76	0.00 7.26	0.00 6.76	0.00 6.26	0.00 5.76	0.00 4.84
5	1.84 11.26	1.47 10.76	0.97 10.26	0.00 9.76	0.00 9.26	0.00 8.76	0.00 8.26	0.00 7.76	0.00 7.26	0.00 6.26
6	2.21 12.75	1.90 12.25	1.61 11.75	1.11 11.25	0.61 10.75	0.11 10.25	0.00 9.75	0.00 9.25	0.00 8.75	0.00 7.75
7	2.58 13.81	2.27 13.31	1.97 12.81	1.69 12.31	1.29 11.81	0.79 11.31	0.29 10.81	0.00 10.31	0.00 9.81	0.00 8.81
8	2.94 15.29	2.63 14.79	2.33 14.29	2.05 13.79	1.78 13.29	1.48 12.79	0.98 12.29	0.48 11.79	0.00 11.29	0.00 10.29
9	4.36 16.77	3.86 16.27	3.36 15.77	2.91 15.27	2.46 14.77	1.96 14.27	1.62 13.77	1.20 13.27	0.70 12.77	0.00 11.77
10	4.75 17.82	4.25 17.32	3.75 16.82	3.30 16.32	2.92 15.82	2.57 15.32	2.25 14.82	1.82 14.32	1.43 13.82	0.43 12.82
11	5.14 19.29	4.64 18.79	4.14 18.29	3.69 17.79	3.30 17.29	2.95 16.79	2.63 16.29	2.33 15.79	2.04 15.29	1.17 14.29
12	6.32 20.34	5.82 19.84	5.32 19.34	4.82 18.84	4.32 18.34	3.85 17.84	3.44 17.34	3.06 16.84	2.69 16.34	1.88 15.34
13	6.72 21.80	6.22 21.30	5.72 20.80	5.22 20.30	4.72 19.80	4.25 19.30	3.84 18.80	3.46 18.30	3.11 17.80	2.47 16.80
14	7.84 22.94	7.34 22.44	6.84 21.94	6.34 21.44	5.84 20.94	5.34 20.44	4.84 19.94	4.37 19.44	3.94 18.94	3.10 17.94
15	8.25 24.31	7.75 23.81	7.25 23.31	6.75 22.81	6.25 22.31	5.75 21.81	5.25 21.31	4.78 20.81	4.35 20.31	3.58 19.31
16	9.34 25.40	8.84 24.90	8.34 24.40	7.84 23.90	7.34 23.40	6.84 22.90	6.34 22.40	5.84 21.90	5.34 21.40	4.43 20.40
17	9.76 26.81	9.26 26.31	8.76 25.81	8.26 25.31	7.76 24.81	7.26 24.31	6.76 23.81	6.26 23.31	5.76 22.81	4.84 21.81
18	10.84 27.84	10.34 27.34	9.84 26.84	9.34 26.34	8.84 25.84	8.34 25.34	7.84 24.84	7.34 24.34	6.84 23.84	5.84 22.84
19	11.26 29.31	10.76 28.81	10.26 28.31	9.76 27.81	9.26 27.31	8.76 26.81	8.26 26.31	7.76 25.81	7.26 25.31	6.26 24.31
20	12.33 30.33	11.83 29.83	11.33 29.33	10.83 28.33	10.33 28.33	9.83 27.83	9.33 27.33	8.83 26.83	8.33 26.33	7.33 25.33

续表

n_0 \ b	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0
0	0.00 1.52	0.00 1.51	0.00 1.50	0.00 1.49	0.00 1.49	0.00 1.48	0.00 1.48	0.00 1.48	0.00 1.47	0.00 1.47
1	0.00 1.78	0.00 1.73	0.00 1.69	0.00 1.66	0.00 1.64	0.00 1.61	0.00 1.60	0.00 1.59	0.00 1.58	0.00 1.56
2	0.00 2.28	0.00 2.11	0.00 1.98	0.00 1.86	0.00 1.81	0.00 1.77	0.00 1.74	0.00 1.72	0.00 1.70	0.00 1.67
3	0.00 2.91	0.00 2.69	0.00 2.37	0.00 2.17	0.00 2.06	0.00 1.98	0.00 1.93	0.00 1.89	0.00 1.82	0.00 1.80
4	0.00 4.05	0.00 3.35	0.00 3.01	0.00 2.54	0.00 2.37	0.00 2.23	0.00 2.11	0.00 2.04	0.00 1.99	0.00 1.95
5	0.00 5.33	0.00 4.52	0.00 3.79	0.00 3.15	0.00 2.94	0.00 2.65	0.00 2.43	0.00 2.30	0.00 2.20	0.00 2.13
6	0.00 6.75	0.00 5.82	0.00 4.99	0.00 4.24	0.00 3.57	0.00 3.14	0.00 2.78	0.00 2.62	0.00 2.48	0.00 2.35
7	0.00 7.81	0.00 6.81	0.00 5.87	0.00 5.03	0.00 4.28	0.00 4.00	0.00 3.37	0.00 3.15	0.00 2.79	0.00 2.59
8	0.00 9.29	0.00 8.29	0.00 7.29	0.00 6.35	0.00 5.50	0.00 4.73	0.00 4.03	0.00 3.79	0.00 3.20	0.00 3.02
9	0.00 10.77	0.00 9.77	0.00 8.77	0.00 7.77	0.00 6.82	0.00 5.96	0.00 5.18	0.00 4.47	0.00 3.81	0.00 3.60
10	0.00 11.82	0.00 10.82	0.00 9.82	0.00 8.82	0.00 7.82	0.00 6.87	0.00 6.00	0.00 5.21	0.00 4.59	0.00 4.24
11	0.17 13.29	0.00 12.29	0.00 11.29	0.00 10.29	0.00 9.29	0.00 8.29	0.00 7.34	0.00 6.47	0.00 5.67	0.00 4.93
12	0.92 14.34	0.00 13.34	0.00 12.34	0.00 11.34	0.00 10.34	0.00 9.34	0.00 8.34	0.00 7.37	0.00 6.50	0.00 5.70
13	1.68 15.80	0.69 14.80	0.00 13.80	0.00 12.80	0.00 11.80	0.00 10.80	0.00 9.80	0.00 8.80	0.00 7.85	0.00 6.96
14	2.28 16.94	1.46 15.94	0.46 14.94	0.00 13.94	0.00 12.94	0.00 11.94	0.00 10.94	0.00 9.94	0.00 8.94	0.00 7.94
15	2.91 18.31	2.11 17.31	1.25 16.31	0.25 15.31	0.00 14.31	0.00 13.31	0.00 12.31	0.00 11.31	0.00 10.31	0.00 9.31
16	3.60 19.40	2.69 18.40	1.98 17.40	1.04 16.40	0.04 15.40	0.00 14.40	0.00 13.40	0.00 12.40	0.00 11.40	0.00 10.40
17	4.05 20.81	3.35 19.81	2.63 18.81	1.83 17.81	0.83 16.81	0.00 15.81	0.00 14.81	0.00 13.81	0.00 12.81	0.00 11.81
18	4.91 21.84	4.11 20.84	3.18 19.84	2.53 18.84	1.63 17.84	0.63 16.84	0.00 15.84	0.00 14.84	0.00 13.84	0.00 12.84
19	5.33 23.31	4.52 22.31	3.79 21.31	3.15 20.31	2.37 19.31	1.44 18.31	0.44 17.31	0.00 16.31	0.00 15.31	0.00 14.31
20	6.33 24.33	5.39 23.33	4.57 22.33	3.82 21.33	2.94 20.33	2.23 19.33	1.25 18.33	0.25 17.33	0.00 16.33	0.00 15.33

表 10.4 置信水平 99%

n_0 \ b	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0
0	0.00 4.74	0.00 4.24	0.00 3.80	0.00 3.50	0.00 3.26	0.00 3.26	0.00 3.05	0.00 3.05	0.00 2.98	0.00 2.94
1	0.01 6.91	0.00 6.41	0.00 5.91	0.00 5.41	0.00 4.91	0.00 4.48	0.00 4.14	0.00 4.09	0.00 3.89	0.00 3.59
2	0.15 8.71	0.00 8.21	0.00 7.71	0.00 7.21	0.00 6.71	0.00 6.24	0.00 5.82	0.00 5.42	0.00 5.06	0.00 4.37
3	0.44 10.47	0.00 9.97	0.00 9.47	0.00 8.97	0.00 8.47	0.00 7.97	0.00 7.47	0.00 6.97	0.00 6.47	0.00 5.57
4	0.82 12.23	0.32 11.73	0.00 11.23	0.00 10.73	0.00 10.23	0.00 9.73	0.00 9.23	0.00 8.73	0.00 8.23	0.00 7.30
5	1.28 13.75	0.78 13.25	0.28 12.75	0.00 12.25	0.00 11.75	0.00 11.25	0.00 10.75	0.00 10.25	0.00 9.75	0.00 8.75
6	1.79 15.27	1.29 14.77	0.79 14.27	0.29 13.77	0.00 13.27	0.00 12.77	0.00 12.27	0.00 11.77	0.00 11.27	0.00 10.27
7	2.33 16.77	1.83 16.27	1.33 15.77	0.83 15.27	0.33 14.77	0.00 14.27	0.00 13.77	0.00 13.27	0.00 12.77	0.00 11.77
8	2.91 18.27	2.41 17.77	1.91 17.27	1.41 16.77	0.91 16.27	0.41 15.77	0.00 15.27	0.00 14.77	0.00 14.27	0.00 13.27
9	3.31 19.46	3.00 18.96	2.51 18.46	2.01 17.96	1.51 17.46	1.01 16.96	0.51 16.46	0.01 15.96	0.00 15.46	0.00 14.46
10	3.68 20.83	3.37 20.33	3.07 19.83	2.63 19.33	2.13 18.83	1.63 18.33	1.13 17.83	0.63 17.33	0.13 16.83	0.00 15.83
11	4.05 22.31	3.73 21.81	3.43 21.31	3.14 20.81	2.77 20.31	2.27 19.81	1.77 19.31	1.27 18.81	0.77 18.31	0.00 17.31
12	4.41 23.80	4.10 23.30	3.80 22.80	3.50 22.30	3.22 21.80	2.93 21.30	2.43 20.80	1.93 20.30	1.43 19.80	0.43 18.80
13	5.83 24.92	5.33 24.42	4.83 23.92	4.33 23.42	3.83 22.92	3.33 22.42	3.02 21.92	2.60 21.42	2.10 20.92	1.10 19.92
14	6.31 26.33	5.81 25.83	5.31 25.33	4.86 24.83	4.46 24.33	4.10 23.83	3.67 23.33	3.17 22.83	2.78 22.33	1.78 21.33
15	6.70 27.81	6.20 27.31	5.70 26.81	5.24 26.31	4.84 25.81	4.48 25.31	4.14 24.81	3.82 24.31	3.42 23.81	2.48 22.81
16	7.76 28.85	7.26 28.35	6.76 27.85	6.26 27.35	5.76 26.85	5.26 26.35	4.76 25.85	4.26 25.35	3.89 24.85	3.15 23.85
17	8.32 30.33	7.82 29.83	7.32 29.33	6.82 28.83	6.32 28.33	5.85 27.83	5.42 27.33	5.03 26.83	4.67 26.33	3.73 25.33
18	8.71 31.81	8.21 31.31	7.71 30.81	7.21 30.31	6.71 29.81	6.24 29.31	5.82 28.81	5.42 28.31	5.06 27.81	4.37 26.81
19	9.88 32.85	9.38 32.35	8.88 31.85	8.38 31.35	7.88 30.85	7.38 30.35	6.88 29.85	6.40 29.35	5.97 28.85	5.01 27.85
20	10.28 34.32	9.78 33.82	9.28 33.32	8.78 32.82	8.28 32.32	7.78 31.82	7.28 31.32	6.81 30.82	6.37 30.32	5.57 29.32

续表

n_0 \ b	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0
0	0.00 2.91	0.00 2.90	0.00 2.89	0.00 2.88	0.00 2.88	0.00 2.87	0.00 2.87	0.00 2.86	0.00 2.86	0.00 2.86
1	0.00 3.42	0.00 3.31	0.00 3.21	0.00 3.18	0.00 3.15	0.00 3.11	0.00 3.09	0.00 3.07	0.00 3.06	0.00 3.03
2	0.00 4.13	0.00 3.89	0.00 3.70	0.00 3.56	0.00 3.44	0.00 3.39	0.00 3.35	0.00 3.32	0.00 3.26	0.00 3.23
3	0.00 5.25	0.00 4.59	0.00 4.35	0.00 4.06	0.00 3.89	0.00 3.77	0.00 3.65	0.00 3.56	0.00 3.51	0.00 3.47
4	0.00 6.47	0.00 5.73	0.00 5.04	0.00 4.79	0.00 4.39	0.00 4.17	0.00 4.02	0.00 3.91	0.00 3.82	0.00 3.74
5	0.00 7.81	0.00 6.97	0.00 6.21	0.00 5.50	0.00 5.17	0.00 4.67	0.00 4.42	0.00 4.24	0.00 4.11	0.00 4.01
6	0.00 9.27	0.00 8.32	0.00 7.47	0.00 6.68	0.00 5.96	0.00 5.46	0.00 5.05	0.00 4.83	0.00 4.63	0.00 4.44
7	0.00 10.77	0.00 9.77	0.00 8.82	0.00 7.95	0.00 7.16	0.00 6.42	0.00 5.73	0.00 5.48	0.00 5.12	0.00 4.82
8	0.00 12.27	0.00 11.27	0.00 10.27	0.00 9.31	0.00 8.44	0.00 7.63	0.00 6.88	0.00 6.18	0.00 5.83	0.00 5.29
9	0.00 13.46	0.00 12.46	0.00 11.46	0.00 10.46	0.00 9.46	0.00 8.50	0.00 7.69	0.00 7.34	0.00 6.62	0.00 5.95
10	0.00 14.83	0.00 13.83	0.00 12.83	0.00 11.83	0.00 10.83	0.00 9.87	0.00 8.98	0.00 8.16	0.00 7.39	0.00 7.07
11	0.00 16.31	0.00 15.31	0.00 14.31	0.00 13.31	0.00 12.31	0.00 11.31	0.00 10.35	0.00 9.46	0.00 8.63	0.00 7.84
12	0.00 17.80	0.00 16.80	0.00 15.80	0.00 14.80	0.00 13.80	0.00 12.80	0.00 11.80	0.00 10.83	0.00 9.94	0.00 9.09
13	0.10 18.92	0.00 17.92	0.00 16.92	0.00 15.92	0.00 14.92	0.00 13.92	0.00 12.92	0.00 11.92	0.00 10.92	0.00 9.98
14	0.78 20.33	0.00 19.33	0.00 18.33	0.00 17.33	0.00 16.33	0.00 15.33	0.00 14.33	0.00 13.33	0.00 12.33	0.00 11.36
15	1.48 21.81	0.48 20.81	0.00 19.81	0.00 18.81	0.00 17.81	0.00 16.81	0.00 15.81	0.00 14.81	0.00 13.81	0.00 12.81
16	2.18 22.85	1.18 21.85	0.18 20.85	0.00 19.85	0.00 18.85	0.00 17.85	0.00 16.85	0.00 15.85	0.00 14.85	0.00 13.85
17	2.89 24.33	1.89 23.33	0.89 22.33	0.00 21.33	0.00 20.33	0.00 19.33	0.00 18.33	0.00 17.33	0.00 16.33	0.00 15.33
18	3.53 25.81	2.62 24.81	1.62 23.81	0.62 22.81	0.00 21.81	0.00 20.81	0.00 19.81	0.00 18.81	0.00 17.81	0.00 16.81
19	4.13 26.85	3.31 25.85	2.35 24.85	1.35 23.85	0.35 22.85	0.00 21.85	0.00 20.85	0.00 19.85	0.00 18.85	0.00 17.85
20	4.86 28.32	3.93 27.32	3.08 26.32	2.08 25.32	1.08 24.32	0.08 23.32	0.00 22.32	0.00 21.32	0.00 20.32	0.00 19.32

表 11　似然比顺序求和方法求得的正态总体期望值 μ 置信水平 68.27%, 90%, 95%, 99% 的置信区间 $[\mu_1, \mu_2]$. x_0 是总体的实验观测值. 表中所有数字均以正态总体标准偏差 σ 为单位

x_0	68.27%C.L.	90%C.L.	95%C.L.	99%C.L.
−3.0	0.00 0.04	0.00 0.26	0.00 0.42	0.00 0.80
−2.9	0.00 0.04	0.00 0.27	0.00 0·44	0.00 0.82
−2.8	0.00 0.04	0.00 0.28	0.00 0.45	0.00 0.84
−2.7	0.00 0.04	0.00 0.29	0.00 0.47	0.00 0.87
−2.6	0.00 0.05	0.00 0.30	0.00 0.48	0.00 0.89
−2.5	0.00 0.05	0.00 0.32	0.00 0.50	0.00 0.92
−2.4	0.00 0.05	0.00 0.33	0.00 0.52	0.00 0.95
−2.3	0.00 0.05	0.00 0.34	0.00 0.54	0.00 0.99
−2.2	0.00 0.06	0.00 0.36	0.00 0.56	0.00 1.02
−2.1	0.00 0.06	0.00 0.38	0.00 0.59	0.00 1.06
−2.0	0.00 0.07	0.00 0.40	0.00 0.62	0.00 1.10
−1.9	0.00 0.08	0.00 0.43	0.00 0.65	0.00 1.14
−1.8	0.00 0.09	0.00 0.45	0.00 0.68	0.00 1.19
−1.7	0.00 0.10	0.00 0.48	0.00 0.72	0.00 1.24
−1.6	0.00 0.11	0.00 0.52	0.00 0.76	0.00 1.29
−1.5	0.00 0.13	0.00 0.56	0.00 0.81	0.00 1.35
−1.4	0.00 0.15	0.00 0.60	0.00 0.86	0.00 1.41
−1.3	0.00 0.17	0.00 0.64	0.00 0.91	0.00 1.47
−1.2	0.00 0.20	0.00 0.70	0.00 0.97	0.00 1.54
−1.1	0.00 0.23	0.00 0.75	0.00 1.04	0.00 1.61
−1.0	0.00 0.27	0.00 0.81	0.00 1.10	0.00 1.68
−0.9	0.00 0.32	0.00 0.88	0.00 1.17	0.00 1.76
−0.8	0.00 0.37	0.00 0.95	0.00 1.25	0.00 1.84
−0.7	0.00 0.43	0.00 1.02	0.00 1.33	0.00 1.93
−0.6	0.00 0.49	0.00 1.10	0.00 1.41	0.00 2.01
−0.5	0.00 0.56	0.00 1.18	0.00 1.49	0.00 2.10
−0.4	0.00 0.64	0.00 1.27	0.00 1.58	0.00 2.19
−0.3	0.00 0.72	0.00 1.36	0.00 1.67	0.00 2.28
−0.2	0.00 0.81	0.00 1.45	0.00 1.77	0.00 2.38
−0.1	0.00 0.90	0.00 1.55	0.00 1.86	0.00 2.48
0.0	0.00 1.00	0.00 1.64	0.00 1.96	0.00 2.58
0.1	0.00 1.10	0.00 1.74	0.00 2.06	0.00 2.68
0.2	0.00 1.20	0.00 1.84	0.00 2.16	0.00 2.78
0.3	0.00 1.30	0.00 1.94	0.00 2.26	0.00 2.88
0.4	0.00 1.40	0.00 2.04	0.00 2.36	0.00 2.98
0.5	0.02 1.50	0.00 2.14	0.00 2.46	0.00 3.08
0.6	0.07 1.60	0.00 2.24	0.00 2.56	0.00 3.18

续表

x_0	68.27%C.L.	90%C.L.	95%C.L.	99%C.L.
0.7	0.11 1.70	0.00 2.34	0.00 2.66	0.00 3.28
0.8	0.15 1.80	0.00 2.44	0.00 2.76	0.00 3.38
0.9	0.19 1.90	0.00 2.54	0.00 2.86	0.00 3.48
1.0	0.24 2.00	0.00 2.64	0.00 2.96	0.00 3.58
1.1	0.30 2.10	0.00 2.74	0.00 3.06	0.00 3.68
1.2	0.35 2.20	0.00 2.84	0.00 3.16	0.00 3.78
1.3	0.42 2.30	0.02 2.94	0.00 3.26	0.00 3.88
1.4	0.49 2.40	0.12 3.04	0.00 3.36	0.00 3.98
1.5	0.56 2.50	0.22 3.14	0.00 3.46	0.00 4.08
1.6	0.64 2.60	0.31 3.24	0.00 3.56	0.00 4.18
1.7	0.72 2.70	0.38 3.34	0.06 3.66	0.00 4.28
1.8	0.81 2.80	0.45 3.44	0.16 3.76	0.00 4.38
1.9	0.90 2.90	0.51 3.54	0.26 3.86	0.00 4.48
2.0	1.00 3.00	0.58 3.64	0.35 3.96	0.00 4.58
2.1	1.10 3.10	0.65 3.74	0.45 4.06	0.00 4.68
2.2	1.20 3.20	0.72 3.84	0.53 4.16	0.00 4.78
2.3	1.30 3.30	0.79 3.94	0.61 4.26	0.00 4.88
2.4	1.40 3.40	0.87 4.04	0.69 4.36	0.07 4.98
2.5	1.50 3.50	0.95 4.14	0.76 4.46	0.17 5.08
2.6	1.60 3.60	1.02 4.24	0.84 4.56	0.27 5.18
2.7	1.70 3.70	1.11 4.34	0.91 4.66	0.37 5.28
2.8	1.80 3.80	1.19 4.44	0.99 4.76	0.47 5.38
2.9	1.90 3.90	1.28 4.54	1.06 4.86	0.57 5.48
3.0	2.00 4.00	1.37 4.64	1.14 4.96	0.67 5.58
3.1	2.10 4.10	1.46 4.74	1.22 5.06	0.77 5.68

表 12　柯尔莫哥洛夫检验临界值 $D_{n,\alpha}^{(+)}$ 表

$$P\left\{D_n^{(+)} > D_{n,\,\alpha}^{(+)}\right\} \leqslant \alpha$$

$D_{n,\alpha}^{+}$	$\alpha=0.10$	0.05	0.025	0.01	0.005
$D_{n,\alpha}$	$\alpha=0.20$	0.10	0.05	0.02	0.01
n=1	0.90000	0.95000	0.97500	0.99000	0.99500
2	0.68377	0.77639	0.84189	0.90000	0.92929
3	0.56481	0.63604	0.70760	0.78456	0.82900
4	0.49265	0.56522	0.62394	0.68887	0.73424
5	0.44698	0.50945	0.56328	0.62718	0.66853
6	0.41037	0.46799	0.51926	0.57741	0.61661
7	0.38148	0.43607	0.48342	0.53844	0.57581
8	0.35831	0.40962	0.45427	0.50654	0.54179
9	0.33910	0.38746	0.43001	0.47960	0.51332

续表

$D^+_{n,\alpha}$	$\alpha=0.10$	0.05	0.025	0.01	0.005
$D_{n,\alpha}$	$\alpha=0.20$	0.10	0.05	0.02	0.01
10	0.32260	0.36866	0.40925	0.45662	0.48893
11	0.30829	0.35242	0.39122	0.43670	0.46770
12	0.29577	0.33815	0.37543	0.41918	0.44905
13	0.28470	0.32549	0.36143	0.40362	0.43247
14	0.27481	0.31417	0.34890	0.38970	0.41762
15	0.26588	0.30397	0.33760	0.37713	0.40420
16	0.25778	0.29472	0.32733	0.36571	0.39201
17	0.25039	0.28627	0.31796	0.35528	0.38086
18	0.24360	0.27851	0.30936	0.34569	0.37062
19	0.23735	0.27136	0.30143	0.33685	0.36117
20	0.23156	0.26473	0.29408	0.32866	0.35241
21	0.22617	0.25858	0.28724	0.32104	0.34427
22	0.22115	0.25283	0.28087	0.31394	0.33666
23	0.21645	0.24746	0.27490	0.30728	0.32954
24	0.21205	0.24242	0.26931	0.30104	0.32286
25	0.20790	0.23768	0.26404	0.29516	0.31657
26	0.20399	0.23320	0.25907	0.28962	0.31064
27	0.20030	0.22898	0.25438	0.28438	0.30502
28	0.19680	0.22497	0.24993	0.27942	0.29971
29	0.19348	0.22117	0.24571	0.27471	0.29466
30	0.19032	0.21756	0.24170	0.27023	0.28987
31	0.18732	0.21412	0.23788	0.26596	0.28530
32	0.18445	0.21085	0.23424	0.26189	0.28094
33	0.18171	0.20771	0.23076	0.25801	0.27677
34	0.17909	0.20472	0.22743	0.25429	0.27279
35	0.17659	0.20185	0.22425	0.25073	0.26897
36	0.17418	0.19910	0.22119	0.24732	0.26532
37	0.17188	0.19646	0.21826	0.24404	0.26180
38	0.16966	0.19392	0.21544	0.24089	0.25843
39	0.16753	0.19148	0.21273	0.23786	0.25518
40	0.16547	0.18913	0.21012	0.23494	0.25205
41	0.16349	0.18687	0.20760	0.23213	0.24904

续表

$D^+_{n,\alpha}$	$\alpha = 0.10$	0.05	0.025	0.01	0.005
$D_{n,\alpha}$	$\alpha = 0.20$	0.10	0.05	0.02	0.01
42	0.16158	0.18468	0.20517	0.22941	0.24613
43	0.15974	0.18257	0.20283	0.22679	0.24332
44	0.15796	0.18053	0.20056	0.22426	0.24060
45	0.15623	0.17856	0.19837	0.22181	0.23798
46	0.15457	0.17665	0.19625	0.21944	0.23544
47	0.15295	0.17481	0.19420	0.21115	0.23298
48	0.15139	0.17302	0.19221	0.21493	0.23059
49	0.14987	0.17128	0.19028	0.21277	0.22828
50	0.14840	0.16959	0.18841	0.21068	0.22604
51	0.14697	0.16796	0.18659	0.20864	0.22386
52	0.14558	0.16637	0.18482	0.20667	0.22174
53	0.14423	0.16483	0.18311	0.20475	0.21968
54	0.14292	0.16332	0.18144	0.20289	0.21768
55	0.14164	0.16186	0.17981	0.20107	0.21574
56	0.14040	0.16044	0.17823	0.19930	0.21384
57	0.13919	0.15906	0.17669	0.19758	0.21199
58	0.13801	0.15771	0.17519	0.19590	0.21019
59	0.13686	0.15639	0.17373	0.19427	0.20844
60	0.13573	0.15511	0.17231	0.19267	0.20673
61	0.13464	0.15385	0.17091	0.19112	0.20506
62	0.13357	0.15263	0.16956	0.18960	0.20343
63	0.13253	0.15144	0.16823	0.18812	0.20184
64	0.13151	0.15027	0.16693	0.18667	0.20029
65	0.13052	0.14913	0.16567	0.18525	0.19877
66	0.12954	0.14802	0.16443	0.18387	0.19729
67	0.12859	0.14693	0.16322	0.18252	0.19584
68	0.12766	0.14587	0.16204	0.18119	0.19442
69	0.12675	0.14483	0.16088	0.17990	0.19303
70	0.12586	0.14381	0.15975	0.17863	0.19167
71	0.12499	0.14281	0.15864	0.17739	0.19034
72	0.12413	0.14183	0.15755	0.17618	0.18903
73	0.12329	0.14087	0.15649	0.17498	0.18776
74	0.12247	0.13993	0.15544	0.17382	0.18650

续表

$D_{n,\alpha}^{+}$	$\alpha=0.10$	0.05	0.025	0.01	0.005
$D_{n,\alpha}$	$\alpha=0.20$	0.10	0.05	0.02	0.01
75	0.12167	0.13901	0.15442	0.17268	0.18528
76	0.12088	0.13811	0.15342	0.17155	0.18408
77	0.12011	0.13723	0.15244	0.17045	0.18290
78	0.11935	0.13636	0.15147	0.16938	0.18174
79	0.11860	0.13551	0.15052	0.16832	0.18060
80	0.11787	0.13467	0.14960	0.16728	0.17949
81	0.11716	0.13385	0.14868	0.16626	0.17840
82	0.11645	0.13305	0.14779	0.16526	0.17732
83	0.11576	0.13226	0.14691	0.16428	0.17627
84	0.11508	0.13148	0.14605	0.16331	0.17523
85	0.11442	0.13072	0.14520	0.16236	0.17421
86	0.11376	0.12997	0.14437	0.16143	0.17321
87	0.11311	0.12923	0.14355	0.16051	0.17223
88	0.11248	0.12850	0.14274	0.15961	0.17126
89	0.11186	0.12779	0.14195	0.15873	0.17031
90	0.11125	0.12709	0.14117	0.15786	0.16938
91	0.11064	0.12640	0.14040	0.15700	0.16846
92	0.11005	0.12572	0.13965	0.15616	0.16755
93	0.10947	0.12506	0.13891	0.15533	0.16666
94	0.10889	0.12440	0.13818	0.15451	0.16579
95	0.10833	0.12375	0.13746	0.15371	0.16493
96	0.10777	0.12312	0.13675	0.15291	0.16408
97	0.10722	0.12249	0.13606	0.15214	0.16324
98	0.10668	0.12187	0.13537	0.15137	0.16242
99	0.10615	0.12126	0.13469	0.15061	0.16161
100	0.10563	0.12067	0.13403	0.14987	0.16081
⩾100	$\frac{1.07}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

表 13　斯米尔诺夫–克拉美–冯　迈希斯检验的显著性水平 α 和统计量 nW^2 的临界值 $(nW^2)_\alpha$ 表

$(nW^2)_\alpha$	α	$(nW^2)_\alpha$	α	$(nW^2)_\alpha$	α	$(nW^2)_\alpha$	α
0.00	1.000000	0.50	0.039833	1.00	0.002460	1.50	0.00173
0.01	0.999994	0.51	0.037575	1.01	0.002331	1.51	0.00164
0.02	0.996999	0.52	0.035451	1.02	0.002209	1.52	0.00155
0.03	0.976168	0.53	0.033453	1.03	0.002093	1.53	0.00148
0.04	0.933149	0.54	0.031573	1.04	0.001983	1.54	0.00140
0.05	0.876281	0.55	0.029803	1.05	0.001880	1.55	0.00133
0.06	0.813980	0.56	0.028136	1.06	0.001781	1.56	0.00126
0.07	0.751564	0.57	0.026566	1.07	0.001688	1.57	0.00120
0.08	0.691855	0.58	0.025088	1.08	0.001600	1.58	0.00114
0.09	0.636144	0.59	0.023695	1.09	0.001516	1.59	0.00108
0.10	0.584873	0.60	0.022382	1.10	0.001437	1.60	0.00102
0.11	0.538041	0.61	0.021145	1.11	0.001362	1.61	0.00097
0.12	0.495425	0.62	0.019978	1.12	0.001291	1.62	0.00092
0.13	0.456707	0.63	0.018878	1.13	0.001224	1.63	0.00087
0.14	0.421539	0.64	0.017841	1.14	0.001160	1.64	0.00083
0.15	0.389576	0.65	0.016862	1.15	0.001100	1.65	0.00079
0.16	0.360493	0.66	0.015939	1.16	0.001043	1.66	0.00075
0.17	0.333995	0.67	0.015068	1.17	0.000989	1.67	0.00071
0.18	0.309814	0.68	0.014246	1.18	0.000937	1.68	0.00067
0.19	0.287709	0.69	0.013470	1.19	0.000889	1.69	0.00064
0.20	0.267470	0.70	0.012738	1.20	0.000843	1.70	0.00061
0.21	0.248908	0.71	0.012046	1.21	0.000799	1.71	0.00058
0.22	0.231856	0.72	0.011393	1.22	0.000758	1.72	0.00055
0.23	0.216167	0.73	0.010776	1.23	0.000718	1.73	0.00052
0.24	0.201710	0.74	0.010194	1.24	0.000681	1.74	0.00049
0.25	0.188370	0.75	0.009644	1.25	0.000646	1.75	0.00047
0.26	0.176042	0.76	0.009124	1.26	0.000613	1.76	0.00044
0.27	0.164636	0.77	0.008633	1.27	0.000581	1.77	0.00042
0.28	0.154070	0.78	0.008169	1.28	0.000551	1.78	0.00040
0.29	0.144270	0.79	0.007730	1.29	0.000522	1.79	0.00038
0.30	0.135171	0.80	0.007316	1.30	0.000496	1.80	0.00036
0.31	0.126715	0.81	0.006924	1.31	0.000470	1.81	0.00034
0.32	0.118847	0.82	0.006554	1.32	0.000446	1.82	0.00032
0.33	0.111522	0.83	0.006203	1.33	0.000423	1.83	0.00031
0.34	0.104695	0.84	0.005872	1.34	0.000401	1.84	0.00029
0.35	0.098327	0.85	0.005559	1.35	0.000380	1.85	0.00028
0.36	0.092383	0.86	0.005263	1.36	0.000361	1.86	0.00026
0.37	0.086832	0.87	0.004983	1.37	0.000342	1.87	0.00025
0.38	0.081642	0.88	0.004718	1.38	0.000325	1.88	0.00024
0.39	0.076789	0.89	0.004468	1.39	0.000308	1.89	0.00023
0.40	0.072247	0.90	0.004231	1.40	0.000292	1.90	0.00021
0.41	0.067995	0.91	0.004007	1.41	0.000277	1.91	0.00020
0.42	0.064010	0.92	0.003795	1.42	0.000263	1.92	0.00019
0.43	0.060276	0.93	0.003594	1.43	0.000249	1.93	0.00018
0.44	0.056774	0.94	0.003404	1.44	0.00237	1.94	0.00017
0.45	0.053488	0.95	0.003225	1.45	0.00225	1.95	0.00017
0.46	0.050405	0.96	0.003054	1.46	0.00213	1.96	0.00016
0.47	0.047510	0.97	0.002893	1.47	0.00202	1.97	0.00015
0.48	0.044790	0.98	0.002741	1.48	0.00192	1.98	0.00014
0.49	0.042235	0.99	0.002597	1.49	0.00182	1.99	0.00013

表 14 Spearman 秩相关检验临界值表

	单侧检验显著性水平			
	0.05	0.025	0.01	0.005
	双侧检验显著性水平			
	0.10	0.05	0.02	0.01
n				
4	1.000	–	–	–
5	0.900	1.000	1.000	–
6	0.829	0.886	0.943	1.000
7	0.714	0.786	0.893	0.929
8	0.643	0.738	0.833	0.881
9	0.600	0.700	0.783	0.833
10	0.564	0.648	0.745	0.794
11	0.536	0.618	0.709	0.755
12	0.503	0.587	0.671	0.727
13	0.484	0.560	0.648	0.703
14	0.464	0.538	0.622	0.675
15	0.443	0.521	0.604	0.654
16	0.429	0.503	0.582	0.635
17	0.414	0.485	0.566	0.615
18	0.401	0.472	0.550	0.600
19	0.391	0.460	0.535	0.584
20	0.380	0.447	0.520	0.570
21	0.370	0.435	0.508	0.556
22	0.361	0.425	0.496	0.544
23	0.353	0.415	0.486	0.532
24	0.344	0.406	0.476	0.521
25	0.337	0.398	0.466	0.511
26	0.331	0.390	0.457	0.501
27	0.324	0.382	0.448	0.491
28	0.317	0.375	0.440	0.483
29	0.312	0.368	0.433	0.475
30	0.306	0.362	0.425	0.467
35	0.283	0.335	0.394	0.433
40	0.264	0.313	0.368	0.405
45	0.248	0.294	0.347	0.382
50	0.235	0.279	0.329	0.363
60	0.214	0.255	0.300	0.331
70	0.190	0.235	0.278	0.307
80	0.185	0.220	0.260	0.287
90	0.174	0.207	0.245	0.271
100	0.165	0.197	0.233	0.257

表 15　Kandall 秩相关检验临界值表

表中给出了 τ 和 S 的临界值.

双侧检验显著性水平	0.01		0.02		0.05		0.10		0.20	
单侧检验显著性水平	0.005		0.01		0.025		0.05		0.10	
n	S	τ	S	τ	S	τ	S	τ	S	τ
4	8	1.000	8	1.000	8	1.000	6	1.000	6	1.000
5	12	1.000	10	1.000	10	1.000	8	0.800	8	0.800
6	15	1.000	13	0.867	13	0.867	11	0.733	9	0.600
7	19	0.905	17	0.810	15	0.714	13	0.619	11	0.524
8	22	0.786	20	0.714	18	0.643	16	0.571	12	0.429
9	26	0.722	24	0.667	20	0.556	18	0.500	14	0.389
10	29	0.644	27	0.600	23	0.511	21	0.467	17	0.378
11	33	0.600	31	0.564	27	0.491	23	0.418	19	0.345
12	38	0.576	36	0.545	30	0.455	26	0.394	20	0.303
13	44	0.564	40	0.513	34	0.436	28	0.359	24	0.308
14	47	0.516	43	0.473	37	0.407	33	0.363	25	0.275
15	53	0.505	49	0.467	41	0.390	35	0.333	29	0.276
16	58	0.483	52	0.433	46	0.383	38	0.317	30	0.250
17	64	0.471	58	0.426	50	0.368	42	0.309	34	0.250
18	69	0.451	63	0.412	53	0.346	45	0.294	37	0.242
19	75	0.439	67	0.392	57	0.333	49	0.287	39	0.228
20	80	0.421	72	0.379	62	0.326	52	0.274	42	0.221
21	86	0.410	78	0.371	66	0.314	56	0.267	44	0.210
22	91	0.394	83	0.359	71	0.307	61	0.264	47	0.203
23	99	0.391	89	0.352	75	0.296	65	0.257	51	0.202
24	104	0.377	94	0.341	80	0.290	68	0.246	54	0.196
25	110	0.367	100	0.333	86	0.287	72	0.240	58	0.193
26	117	0.360	107	0.329	91	0.280	77	0.237	61	0.188
27	125	0.356	113	0.322	95	0.271	81	0.231	63	0.179
28	130	0.344	118	0.312	100	0.265	86	0.228	68	0.180
29	138	0.340	126	0.310	106	0.261	90	0.222	70	0.172
30	145	0.333	131	0.301	111	0.255	95	0.218	75	0.172
31	151	0.325	137	0.295	117	0.252	99	0.213	77	0.166
32	160	0.323	144	0.290	122	0.246	104	0.210	82	0.165
33	166	0.314	152	0.288	128	0.242	108	0.205	86	0.163
34	175	0.312	157	0.280	133	0.237	113	0.201	89	0.159
35	181	0.304	165	0.277	139	0.234	117	0.197	93	0.156
36	190	0.302	172	0.273	146	0.232	122	0.194	96	0.152
37	198	0.297	178	0.267	152	0.228	128	0.192	100	0.150
38	205	0.292	185	0.263	157	0.223	133	0.189	105	0.149
39	213	0.287	193	0.260	163	0.220	139	0.188	109	0.147
40	222	0.285	200	0.256	170	0.218	144	0.185	112	0.144

表 16 多维变量 Kandall 协和系数 *W* 的临界值表

表中给出了 W 和 S 的临界值. (n 为样本数, m 为变量维数)

m	n=3		4		5		6		7	
	显著性 $\alpha=0.05$									
	S	$\bar{W}$	S	$\bar{W}$	S	$\bar{W}$	S	$\bar{W}$	S	$\bar{W}$
3					64.4	0.716	103.9	0.660	157.3	0.624
4			49.5	0.619	88.4	0.522	143.3	0.512	217.0	0.484
5			62.6	0.501	112.3	0.449	182.4	0.417	276.2	0.395
6			75.7	0.421	136.1	0.378	221.4	0.351	335.2	0.333
8	48.1	0.376	101.7	0.318	183.7	0.287	299.0	0.267	453.1	0.253
10	60.0	0.300	127.8	0.256	231.2	0.231	376.7	0.215	571.0	0.204
15	89.9	0.200	192.9	0.171	349.8	0.155	570.5	0.145	864.9	0.137
20	119.7	0.150	258.0	0.129	468.5	0.117	764.4	0.109	1158.7	0.103
	显著性 $\alpha=0.01$									
3					75.6	0.840	122.8	0.780	185.6	0.737
4			61.4	0.768	109.3	0.683	176.2	0.629	265.0	0.592
5			80.5	0.644	142.8	0.571	229.4	0.524	343.8	0.491
6			99.5	0.553	176.1	0.489	282.4	0.448	422.6	0.419
8	66.8	0.522	137.4	0.429	242.7	0.379	388.3	0.347	579.9	0.324
10	85.1	0.425	175.3	0.351	309.1	0.309	494.0	0.282	737.0	0.263
15	131.0	0.291	269.8	0.240	475.2	0.211	758.2	0.193	1129.5	0.179
20	177.0	0.221	364.2	0.182	641.2	0.160	1022.2	0.146	1521.9	0.136

$n=3$				
	$\alpha=0.05$		$\alpha=0.01$	
m	S	$\bar{W}$	S	$\bar{W}$
9	54.0	0.333	75.9	0.469
12	71.9	0.250	103.5	0.359
14	83.8	0.214	121.9	0.311
16	95.8	0.187	140.2	0.274
18	107.7	0.166	158.6	0.245

表 17　游程检验的临界值 R_α 表

$$\sum_{R=2}^{R_\alpha} p(R) \leqslant \alpha < \sum_{R=2}^{R_\alpha+1} p(R)$$

$R_{0.025}$

m \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4																			
5			2	2															
6		2	2	3	3														
7		2	2	3	3	3													
8		2	3	3	3	4	4												
9		2	3	3	4	4	5	5											
10		2	3	3	4	5	5	5	6										
11		2	3	4	4	5	5	6	6	7									
12	2	2	3	4	4	5	6	6	7	7	7								
13	2	2	3	4	5	5	6	6	7	7	8	8							
14	2	2	3	4	5	5	6	7	7	8	8	9	9						
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10					
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11				
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11			
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12		
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14

$R_{0.005}$

m \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4																			
5																			
6				2	2														
7				2	2	3													
8			2	2	3	3	3												
9			2	2	3	3	3	4											
10			2	3	3	3	4	4	5										
11			2	3	3	4	4	5	5	5									
12		2	2	3	3	4	4	5	5	6	6								
13		2	2	3	3	4	5	5	5	6	6	7							
14		2	2	3	4	4	5	5	6	6	7	7	7						
15		2	3	3	4	4	5	6	6	7	7	7	8	8					
16		2	3	3	4	5	5	6	6	7	7	8	8	9	9				
17		2	3	3	4	5	5	6	7	7	8	8	8	9	9	10			
18		2	3	4	4	5	6	6	7	7	8	8	9	9	10	10	11		
19		2	3	4	4	5	6	6	7	8	8	9	9	10	10	10	11	11	
20		2	3	4	4	5	6	7	7	8	8	9	9	10	10	11	11	11	12

续表

$R_{0.05}$

m \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4			2																
5		2	2	3															
6		2	3	3	3														
7		2	2	3	4	4													
8	2	2	3	3	4	4	5												
9	2	2	3	4	4	5	5	6											
10	2	3	3	4	5	5	6	6	6										
11	2	3	3	4	5	5	6	6	7	7									
12	2	3	4	4	5	6	6	7	7	8	8								
13	2	3	4	4	5	6	6	7	8	8	9	9							
14	2	3	4	5	5	6	7	7	8	8	9	9	10						
15	2	3	4	5	6	6	7	8	8	9	9	10	10	11					
16	2	3	4	5	6	6	7	8	8	9	10	10	11	11	11				
17	2	3	4	5	6	7	7	8	9	9	10	10	11	11	12	12			
18	2	3	4	5	6	7	8	8	9	10	10	11	11	12	12	13	13		
19	2	3	4	5	6	7	8	8	9	10	10	11	12	12	13	13	14	14	
20	2	3	4	5	6	7	8	9	9	10	11	11	12	12	13	13	14	14	15

$R_{0.010}$

m \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4																			
5				2															
6			2	2	2														
7			2	2	3	3													
8			2	2	3	3	4												
9		2	2	3	3	4	4	4											
10		2	2	3	3	4	4	5	5										
11		2	2	3	4	4	5	5	5	6									
12		2	3	3	4	4	5	5	6	6	7								
13		2	3	3	4	5	5	6	6	6	7	7							
14		2	3	3	4	5	5	6	6	7	7	8	8						
15		2	3	4	4	5	5	6	7	7	8	8	8	9					
16		2	3	4	4	5	6	6	7	7	8	8	9	9	10				
17		2	3	4	5	5	6	7	7	8	8	9	9	10	10	10			
18		2	3	4	5	5	6	7	7	8	8	9	9	10	10	11	11		
19	2	2	3	4	5	6	6	7	8	8	9	9	10	10	11	11	12	12	
20	2	2	3	4	5	6	6	7	8	8	9	10	10	11	11	11	12	12	13

表 18　斯米尔诺夫检验临界值 $D_{n,n;\alpha}^{(+)}$ 表

$$P\left\{D_{n,n}^{(+)} > D_{n,n;\alpha}^{(+)}\right\} \leqslant \alpha$$

$D_{n,n;\alpha}^{+}$	$\alpha = 0.10$	0.05	0.025	0.01	0.005
$D_{n,n;\alpha}$	$\alpha = 0.20$	0.10	0.05	0.02	0.01
$n = 3$	2/3	2/3			
4	3/4	3/4	3/4		
5	3/5	3/5	4/5	4/5	4/5
6	3/6	4/6	4/6	5/6	5/6
7	4/7	4/7	5/7	5/7	5/7
8	4/8	4/8	5/8	5/8	6/8
9	4/9	5/9	5/9	6/9	6/9
10	4/10	5/10	6/10	6/10	7/10
11	5/11	5/11	6/11	7/11	7/11
12	5/12	5/12	6/12	7/12	7/12
13	5/13	6/13	6/13	7/13	8/13
14	5/14	6/14	7/14	7/14	8/14
15	5/15	6/15	7/15	8/15	8/15
16	6/16	6/16	7/16	8/16	9/16
17	6/17	7/17	7/17	8/17	9/17
18	6/18	7/18	8/18	9/18	9/18
19	6/19	7/19	8/19	9/19	9/19
20	6/20	7/20	8/20	9/20	10/20
21	6/21	7/21	8/21	9/21	10/21
22	7/22	8/22	8/22	10/22	10/22
23	7/23	8/23	9/23	10/23	10/23
24	7/24	8/24	9/24	10/24	11/24
25	7/25	8/25	9/25	10/25	11/25
26	7/26	8/26	9/26	10/26	11/26
27	7/27	8/27	9/27	11/27	11/27
28	8/28	9/28	10/28	11/28	12/28
29	8/29	9/29	10/29	11/29	12/29
30	8/30	9/30	10/30	11/30	12/30
31	8/31	9/31	10/31	11/31	12/31
32	8/32	9/32	10/32	12/32	12/32
34	8/34	10/34	11/34	12/34	13/34
36	9/36	10/36	11/36	12/36	13/36
38	9/38	10/38	11/38	13/38	14/38
40	9/40	10/40	12/40	13/40	14/40
对 $n > 40$ 的近似	$\frac{1.52}{\sqrt{n}}$	$\frac{1.73}{\sqrt{n}}$	$\frac{1.92}{\sqrt{n}}$	$\frac{2.15}{\sqrt{n}}$	$\frac{2.30}{\sqrt{n}}$

表 19　斯米尔诺夫检验临界值 $D^{(+)}_{m,n;\alpha}$ 表

$$P\left\{D^{(+)}_{m,n} > D^{(+)}_{m,n;\alpha}\right\} \leqslant \alpha$$

表中 $N_1 = \min(x, n)$, $N_2 = \max(m, n)$

$D^{+}_{m,n;\alpha}$		$\alpha = 0.10$	0.05	0.025	0.01	0.005
$D_{m,n;\alpha}$		$\alpha = 0.20$	0.10	0.05	0.02	0.01
N_1=1	N_2=9	17/18				
	10	9/10				
N_1=2	N_2=3	5/6				
	4	3/4				
	5	4/5	4/5			
	6	5/6	5/6			
	7	5/7	6/7			
	8	3/4	7/8	7/8		
	9	7/9	8/9	8/9		
	10	7/10	4/5	9/10		
N_1=3	N_2=4	3/4	3/4			
	5	2/3	4/5	4/5		
	6	2/3	2/3	5/6		
	7	2/3	5/7	6/7	6/7	
	8	5/8	3/4	3/4	7/8	
	9	2/3	2/3	7/9	8/9	8/9
	10	3/5	7/10	4/5	9/10	9/10
	12	7/12	2/3	3/4	5/6	11/12
N_1=4	N_2=5	3/5	3/4	4/5	4/5	
	6	7/12	2/3	3/4	5/6	5/6
	7	17/28	5/7	3/4	6/7	6/7
	8	5/8	5/8	3/4	7/8	7/8
	9	5/9	2/3	3/4	7/9	8/9
	10	11/20	13/20	7/10	4/5	4/5
	12	7/12	2/3	2/3	3/4	5/6
	16	9/16	5/8	11/16	3/4	13/16
N_1=5	N_2=6	3/5	2/3	2/3	5/6	5/6
	7	4/7	23/35	5/7	29/35	6/7
	8	11/20	5/8	27/40	4/5	4/5
	9	5/9	3/5	31/45	7/9	4/5
	10	1/2	3/5	7/10	7/10	4/5
	15	8/15	3/5	2/3	11/15	11/15
	20	1/2	11/20	3/5	7/10	3/4

续表

$D^{+}_{m,n;\alpha}$		$\alpha=0.10$	0.05	0.025	0.01	0.005
$D_{m,n;\alpha}$		$\alpha=0.20$	0.10	0.05	0.02	0.01
N_1=6	N_2=7	23/42	4/7	29/42	5/7	5/6
	8	1/2	7/12	2/3	3/4	3/4
	9	1/2	5/9	2/3	13/18	7/9
	10	1/2	17/30	19/30	7/10	11/15
	12	1/2	7/12	7/12	2/3	3/4
	18	4/9	5/9	11/18	2/3	13/18
	24	11/24	1/2	7/12	5/8	2/3
N_1=7	N_2=8	27/56	33/56	5/8	41/56	3/4
	9	31/63	5/9	40/63	5/7	47/63
	10	33/70	39/70	43/70	7/10	5/7
	14	3/7	1/2	4/7	9/14	5/7
	28	3/7	13/28	15/28	17/28	9/14
N_1=8	N_2=9	4/9	13/24	5/8	2/3	3/4
	10	19/40	21/40	23/40	27/40	7/10
	12	11/24	1/2	7/12	5/8	2/3
	16	7/16	1/2	9/16	5/8	5/8
	32	13/32	7/16	1/2	9/16	19/32
N_1=9	N_2=10	7/15	1/2	26/45	2/3	31/45
	12	4/9	1/2	5/9	11/18	2/3
	15	19/45	22/45	8/15	3/5	29/45
	18	7/18	4/9	1/2	5/9	11/18
	36	13/36	5/12	17/36	19/36	5/9
N_1=10	N_2=15	2/5	7/15	1/2	17/30	19/30
	20	2/5	9/20	1/2	11/20	3/5
	40	7/20	2/5	9/20	1/2	
N_1=12	N_2=15	23/60	9/20	1/2	11/20	7/12
	16	3/8	7/16	23/48	13/24	7/12
	18	13/36	5/12	17/36	19/36	5/9
	20	11/30	5/12	7/15	31/60	17/30
N_1=15	N_2=20	7/20	2/5	13/30	29/60	31/60
N_1=16	N_2=20	27/80	31/80	17/40	19/40	41/80
$m,\ n\to\infty$		$1.07\sqrt{\frac{m+n}{mn}}$	$1.22\sqrt{\frac{m+n}{mn}}$	$1.36\sqrt{\frac{m+n}{mn}}$	$1.52\sqrt{\frac{m+n}{mn}}$	$1.63\sqrt{\frac{m+n}{mn}}$

表 20 威尔科克森秩和检验临界值 W_α 表

$$\sum_{W_{\min}}^{W_a} P(W) \leqslant \alpha < \sum_{W_{\min}}^{W_a+1} P(W)$$

$2 \leqslant m \leqslant 25$, $n \leqslant m$. 六种显著性水平 α 在每个表的第一行以粗体字组出, $\bar{W}$ 的含义见 12.7.5 节.

			$n=1$								$n=2$				
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
2							4						—	10	2
3							5						3	12	3
4							6					—	3	14	4
5							7					3	4	16	5
6							8					3	4	18	6
7							9				—	3	4	20	7
8						—	10				3	4	5	22	8
9						1	11				3	4	5	24	9
10						1	12				3	4	6	26	10
11						1	13				3	4	6	28	11
12						1	14			—	4	5	7	30	12
13						1	15			3	4	5	7	32	13
14						1	16			3	4	6	8	34	14
15						1	17			3	4	6	8	36	15
16						1	18			3	4	6	8	38	16
17						1	19			3	5	6	9	40	17
18					—	1	20		—	3	5	7	9	42	18
19					1	2	21		3	4	5	7	10	44	19
20					1	2	22		3	4	5	7	10	46	20
21					1	2	23		3	4	6	8	11	48	21
22					1	2	24		3	4	6	8	11	50	22
23					1	2	25		3	4	6	8	12	52	23
24					1	2	26		3	4	6	9	12	54	24
25	—	—	—	—	1	2	27		3	4	6	9	12	56	25

			$n=3$								$n=4$				
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
3					6	7	21								
4				—	6	7	24			—	10	11	13	36	4
5				6	7	8	27		—	10	11	12	14	40	5
6			—	7	8	9	30		10	11	12	13	15	44	6
7			6	7	8	10	33		10	11	13	14	16	48	7
8		—	6	8	9	11	36		11	12	14	15	17	52	8

续表

			$n=3$								$n=4$				
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
9		6	7	8	10	11	39	—	11	13	14	16	19	56	9
10		6	7	9	10	12	42	10	12	13	15	17	20	60	10
11		6	7	9	11	13	45	10	12	14	16	18	21	64	11
12		7	8	10	11	14	48	10	13	15	17	19	22	68	12
13		7	8	10	12	15	51	11	13	15	18	20	23	72	13
14		7	8	11	13	16	54	11	14	16	19	21	25	76	14
15		8	9	11	13	16	57	11	15	17	20	22	26	80	15
16	—	8	9	12	14	17	60	12	15	17	21	24	27	84	16
17	6	8	10	12	15	18	63	12	16	18	21	25	28	88	17
18	6	8	10	13	15	19	66	13	16	19	22	26	30	92	18
19	6	9	10	13	16	20	69	13	17	19	23	27	31	96	19
20	6	9	11	14	17	21	72	13	18	20	24	28	32	100	20
21	7	9	11	14	17	21	75	14	18	21	25	29	33	104	21
22	7	10	12	15	18	22	78	14	19	21	26	30	35	108	22
23	7	10	12	15	19	23	81	14	19	22	27	31	36	112	23
24	7	10	12	16	19	24	84	15	20	23	27	32	38	116	24
25	7	11	13	16	20	25	87	15	20	23	28	33	38	120	25

			$n=5$								$n=6$				
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
5		15	16	17	19	20	55								
6		16	17	18	20	22	60	—	23	24	26	28	30	78	6
7	—	16	18	20	21	23	65	21	24	25	27	29	32	84	7
8	15	17	19	21	23	25	70	22	25	27	29	31	34	90	8
9	16	18	20	22	24	27	75	23	26	28	31	33	36	96	9
10	16	19	21	23	26	28	80	24	27	29	32	35	38	102	10
11	17	20	22	24	27	30	85	25	28	30	34	37	40	108	11
12	17	21	23	26	28	32	90	25	30	32	35	38	42	114	12
13	18	22	24	27	30	33	95	26	31	33	37	40	44	120	13
14	18	22	25	28	31	35	100	27	32	34	38	42	46	126	14
15	19	23	26	29	33	37	105	28	33	36	40	44	48	132	15
16	20	24	27	30	34	38	110	29	34	37	42	46	50	138	16
17	20	25	28	32	35	40	115	30	38	39	43	47	52	144	17
18	21	26	29	33	37	42	120	31	37	40	45	49	55	150	18

续表

			$n=5$								$n=6$				
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
19	22	27	30	34	38	43	125	32	38	41	46	51	57	156	19
20	22	28	31	35	40	45	130	33	39	43	48	53	59	162	20
21	23	29	32	37	41	47	135	33	40	44	50	55	61	168	21
22	23	29	33	38	43	48	140	34	42	45	51	57	63	174	22
23	24	30	34	39	44	50	145	35	43	47	53	58	65	180	23
24	25	31	35	40	45	51	150	36	44	48	54	60	67	186	24
25	25	32	36	42	47	53	155	37	45	50	56	62	69	192	25

			$n=7$								$n=8$				
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
7	29	32	34	36	39	41	105								
8	30	34	35	38	41	44	112	40	43	45	49	51	55	136	8
9	31	35	37	40	43	46	119	41	45	47	51	54	58	144	9
10	33	37	39	42	45	49	126	42	47	49	53	56	60	152	10
11	34	38	40	44	47	51	133	44	49	51	55	59	63	160	11
12	35	40	42	46	49	54	140	45	51	53	58	62	66	168	12
13	36	41	44	48	52	56	147	47	53	56	60	64	69	176	13
14	37	43	45	50	54	59	154	48	54	58	62	67	72	184	14
15	38	44	47	52	56	61	161	50	56	60	65	69	75	192	15
16	39	46	49	54	58	64	168	51	58	62	67	72	78	200	16
17	41	47	51	56	61	66	175	53	60	64	70	75	81	208	17
18	42	49	52	58	63	69	182	54	62	66	72	77	84	216	18
19	43	50	54	60	65	71	189	56	64	68	74	80	87	224	19
20	44	52	56	62	67	74	196	57	66	70	77	83	90	232	20
21	46	53	58	64	69	76	203	59	68	72	79	85	92	240	21
22	47	55	59	66	72	79	210	60	70	74	81	88	95	248	22
23	48	57	61	68	74	81	217	62	71	76	84	90	98	256	23
24	49	58	63	70	76	84	224	64	73	78	86	93	101	264	24
25	50	60	64	72	78	86	231	65	75	81	89	96	104	272	25

			$n=9$								$n=10$				
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
9	52	56	59	62	66	70	171								
10	53	58	61	65	69	73	180	65	71	74	78	82	87	210	10
11	55	61	63	68	72	76	189	67	73	77	81	86	91	220	11

续表

			n = 9								n = 10				
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
12	57	63	66	71	75	80	198	69	76	79	84	89	94	230	12
13	59	65	68	73	78	83	207	72	79	82	88	92	98	240	13
14	60	67	71	76	81	86	216	74	81	85	91	96	102	250	14
15	62	69	73	79	84	90	225	76	84	88	94	99	106	260	15
16	64	72	76	82	87	93	234	78	86	91	97	103	109	270	16
17	66	74	78	84	90	97	243	80	89	93	100	106	113	280	17
18	68	76	81	87	93	100	252	82	92	96	103	110	117	290	18
19	70	78	83	90	96	103	261	84	94	99	107	113	121	300	19
20	71	81	85	93	99	107	270	87	97	102	110	117	125	310	20
21	73	83	88	95	102	110	279	89	99	105	113	120	128	320	21
22	75	85	90	98	105	113	288	91	102	108	116	123	132	330	22
23	77	88	93	101	108	117	297	93	105	110	119	127	136	340	23
24	79	90	95	104	111	120	306	95	107	113	122	130	140	350	24
25	81	92	98	107	114	123	315	98	110	116	126	134	144	360	25

			n = 11								n = 12				
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
11	81	87	91	96	100	106	253								
12	83	90	94	99	104	110	264	98	105	109	115	120	127	300	12
13	86	93	97	103	108	114	275	101	109	113	119	125	131	312	13
14	88	96	100	106	112	118	286	103	112	116	123	129	136	324	14
15	90	99	103	110	116	123	297	106	115	120	127	133	141	336	15
16	93	102	107	113	120	127	308	109	119	124	131	138	145	348	16
17	95	105	110	117	123	131	319	112	122	127	135	142	150	360	17
18	98	108	113	121	127	135	330	115	125	131	139	146	155	372	18
19	100	111	116	124	131	139	341	118	129	134	143	150	159	384	19
20	103	114	119	128	135	144	352	120	132	138	147	155	164	396	20
21	106	117	123	131	139	148	363	123	136	142	151	159	169	408	21
22	108	120	126	135	143	152	374	126	139	145	155	163	173	420	22
23	111	123	129	139	147	156	385	129	142	149	159	168	178	432	23
24	113	126	132	142	151	161	396	132	146	153	163	172	183	444	24
25	116	129	136	146	155	165	407	135	149	156	167	176	187	456	25

			n = 13								n = 14				
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
13	117	125	130	136	142	149	351								

续表

	n = 13							n = 14							
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
14	120	129	134	141	147	154	364	137	147	152	160	166	174	406	14
15	123	133	138	145	152	159	377	141	151	156	164	171	179	420	15
16	126	136	142	150	156	165	390	144	155	161	169	176	185	434	16
17	129	140	146	154	161	170	403	148	159	165	174	182	190	448	17
18	133	144	150	158	166	175	416	151	163	170	179	187	196	462	18
19	136	148	154	163	171	180	429	155	168	174	183	192	202	476	19
20	139	151	158	167	175	185	442	159	172	178	188	197	207	490	20
21	142	155	162	171	180	190	455	162	176	183	193	202	213	504	21
22	145	159	166	176	185	195	468	166	180	187	198	207	218	518	22
23	149	163	170	180	189	200	481	169	184	192	203	212	224	532	23
24	152	166	174	185	194	205	494	173	188	196	207	218	229	546	24
25	155	170	178	189	199	211	507	177	192	200	212	223	235	560	25

	n = 15							n = 16							
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
15	160	171	176	184	192	200	465								
16	163	175	181	190	197	206	480	184	196	202	211	219	229	528	16
17	167	180	186	195	203	212	495	188	201	207	217	225	235	544	17
18	171	184	190	200	208	218	510	192	206	212	222	231	242	560	18
19	175	189	195	205	214	224	525	196	210	218	228	237	248	576	19
20	179	193	200	210	220	230	540	201	215	223	234	243	255	592	20
21	183	198	205	216	225	236	555	205	220	228	239	249	261	608	21
22	187	202	210	221	231	242	570	209	225	233	245	255	267	624	22
23	191	207	214	226	236	248	585	214	230	238	251	261	274	640	23
24	195	211	219	231	242	254	600	218	235	244	256	267	280	656	24
25	199	216	224	237	248	260	615	222	240	249	262	273	287	672	25

	n = 17							n = 18							
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
17	210	223	230	240	249	259	595								
18	214	228	235	246	255	266	612	237	252	259	270	280	291	666	18
19	219	234	241	252	262	273	629	242	258	265	277	287	299	684	19
20	223	239	246	258	268	280	646	247	263	271	283	294	306	702	20
21	228	244	252	264	274	287	663	252	269	277	290	301	313	720	21
22	233	249	258	270	281	294	680	257	275	283	296	307	321	738	22

续表

	$n=17$							$n=18$							
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
23	238	255	263	276	287	300	697	262	280	289	303	314	328	756	23
24	242	260	269	282	294	307	714	267	286	295	309	321	335	774	24
25	247	265	275	288	300	314	731	273	292	301	316	328	343	792	25

	$n=19$							$n=20$							
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
19	267	283	291	303	313	325	741								
20	272	289	297	309	320	333	760	298	315	324	337	348	361	820	20
21	277	295	303	316	328	341	779	304	322	331	344	356	370	840	21
22	283	301	310	323	335	349	798	309	328	337	351	364	378	860	22
23	288	307	316	330	342	357	817	315	335	344	359	371	386	880	23
24	294	313	323	337	350	364	836	321	341	351	366	379	394	900	24
25	299	319	329	344	357	372	855	327	348	358	373	387	403	920	25

	$n=21$							$n=22$							
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
21	331	349	359	373	385	399	903								
22	337	356	366	381	393	408	924	365	386	396	411	424	439	990	22
23	343	363	373	388	401	417	945	372	393	403	419	432	448	1012	23
24	349	370	381	396	410	425	966	379	400	411	427	441	457	1034	24
25	356	377	388	404	418	434	987	385	408	419	435	450	467	1056	25

	$n=23$							$n=24$							
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$	m
23	402	424	434	451	465	481	1081								
24	409	431	443	459	474	491	1104	440	464	475	492	507	525	1176	24
25	416	439	451	468	483	500	1127	448	472	484	501	517	535	1200	25

	$n=25$						
m	0.001	0.005	0.010	0.025	0.05	0.10	$2\bar{W}$
25	480	505	517	536	552	570	1275

表 21　威尔科克森符号秩和检验临界值 W_α 表

n	单侧检验 $\alpha = 0.01$	$\alpha = 0.025$	$\alpha = 0.05$
	双侧检验 $\alpha = 0.02$	$\alpha = 0.05$	$\alpha = 0.10$
5			1
6		1	2
7	0	2	4
8	2	4	6
9	3	6	8
10	5	8	11
11	7	11	14
12	10	14	17
13	13	17	21
14	16	21	26
15	20	25	30
16	24	30	36
17	28	35	41
18	33	40	47
19	38	46	54
20	43	52	60
21	49	59	68
22	56	66	75
23	62	73	83
24	69	81	92
25	77	90	101
26	85	98	110
27	93	107	120
28	102	117	130
29	111	127	141
30	120	137	152

表 22　克鲁斯卡尔–瓦列斯检验临界值 H_α 表

子样数 $J = 3$, 子样容量 $n_1, n_2, n_3 \leqslant 5$. $P\{H \geqslant H_\alpha\} = \alpha$

子样容量					子样容量				
n_1	n_2	n_3	H_α	α	n_1	n_2	n_3	H_α	α
2	1	1	2.7000	0.500	4	2	2	6.0000	0.014
								5.3333	0.033
2	2	1	3.6000	0.200				5.1250	0.052
								4.4583	0.100
2	2	2	4.5714	0.067				4.1667	0.105
			3.7143	0.200					
					4	3	1	5.8333	0.021
3	1	1	3.2000	0.300				5.2083	0.050
								5.0000	0.057
3	2	1	4.2857	0.100				4.0556	0.093

续表

子样容量					子样容量				
n_1	n_2	n_3	H_α	α	n_1	n_2	n_3	H_α	α
			3.8571	0.133				3.8889	0.129
3	2	2	5.3572	0.029	4	3	2	6.4444	0.008
			4.7143	0.048				6.3000	0.010
			4.5000	0.067				5.4444	0.046
			4.4643	0.105				5.4000	0.051
								4.5111	0.098
3	3	1	5.1429	0.043				4.4444	0.102
			4.5714	0.100					
			4.0000	0.129	4	3	3	6.7455	0.010
								6.7091	0.013
3	3	2	6.2500	0.011				5.7909	0.046
			5.3611	0.032				5.7203	0.050
			5.1389	0.061				4.7091	0.092
			4.5556	0.100				4.7000	0.101
			4.2500	0.121					
					4	4	1	6.6667	0.010
3	3	3	7.2000	0.004				6.1667	0.022
			6.4889	0.011				4.9667	0.048
			5.6889	0.029				4.8667	0.054
			5.6000	0.050				4.1667	0.082
			5.0667	0.086				4.0667	0.102
			4.6222	0.100					
					4	4	2	7.0364	0.006
4	1	1	3.5714	0.200				6.8727	0.011
								5.4545	0.046
4	2	1	4.8214	0.057				5.2364	0.052
			4.5000	0.076				4.5545	0.098
			4.0179	0.114				4.4455	0.103
4	4	3	7.1439	0.010	5	3	3	7.0788	0.009
			7.1364	0.011				6.9818	0.011
			5.5985	0.049				5.6485	0.049
			5.5758	0.051				5.5152	0.051
			4.5455	0.099				4.5333	0.097
			4.4773	0.102				4.4121	0.109
4	4	4	7.6538	0.008	5	4	1	6.9545	0.008
			7.5385	0.011				6.8400	0.011
			5.6923	0.049				4.9855	0.044
			5.6538	0.054				4.8600	0.056
			4.6539	0.097				3.9873	0.098
			4.5001	0.104				3.9600	0.102
5	1	1	3.8571	0.143	5	4	2	7.2045	0.009

续表

子样容量				
n_1	n_2	n_3	H_α	α
			7.1182	0.010
			5.2727	0.049
			5.2682	0.050
			4.5409	0.098
			4.5182	0.101
5	2	1	5.2500	0.036
			5.0000	0.048
			4.4500	0.071
			4.2000	0.095
			4.0500	0.119
5	2	2	6.5333	0.008
			6.1333	0.013
			5.1600	0.034
			5.0400	0.056
			4.3733	0.090
			4.2933	0.122
5	3	1	6.4000	0.012
			4.9600	0.048
			4.8711	0.052
			4.0178	0.095
			3.8400	0.123
5	3	2	6.9091	0.009
			6.8218	0.010
			5.2509	0.049
			5.1055	0.052
			4.6509	0.091
			4.4945	0.101
5	4	3	7.4449	0.010
			7.3949	0.011
			5.6564	0.049
			5.6308	0.050
			4.5487	0.099
			4.5231	0.103
5	4	4	7.7604	0.009
			7.7440	0.011
			5.6571	0.049
			5.6176	0.050
			4.6187	0.100
			4.5527	0.102
5	5	1	7.3091	0.009
			6.8364	0.011
			5.1273	0.046
			4.9091	0.053
			4.1091	0.086
			4.0364	0.105
5	5	2	7.3385	0.010
			7.2692	0.010
			5.3385	0.047
			5.2462	0.051
			4.6231	0.097
			4.5077	0.100
5	5	3	7.5780	0.010
			7.5429	0.010
			5.7055	0.046
			5.6264	0.051
			4.5451	0.100
			4.5363	0.102
5	5	4	7.8229	0.010
			7.7914	0.010
			5.6657	0.049
			5.6429	0.050
			4.5229	0.099
			4.5200	0.101
5	5	5	8.0000	0.009
			7.9800	0.010
			5.7800	0.049
			5.6600	0.051
			4.5600	0.100
			4.5000	0.102

表 23.1 信度水平 68.27%

$n\backslash b$	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0
0	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15
1	0.27,2.50	0.11,1.99	0.00,1.79	0.00,1.66	1.00,1.57	0.00,1.51	0.00,1.46	0.00,1.42	0.00,1.39	0.00,1.35
2	0.87,3.85	0.38,3.30	0.00,2.66	0.00,2.37	0.00,2.16	0.00,2.01	0.00,1.89	0.00,1.80	0.00,1.72	0.00,1.61
3	1.56,5.14	1.06,4.64	0.58,4.08	0.16,3.42	0.00,2.95	0.00,2.68	0.00,2.47	0.00,2.30	0.00,2.16	0.00,1.96
4	2.29,6.40	1.79,5.90	1.30,5.39	0.83,4.83	0.39,4.20	0.00,3.52	0.00,3.20	0.00,2.94	0.00,2.73	0.00,2.40
5	3.06,7.63	2.56,7.13	2.06,6.63	1.57,6.11	1.10,5.56	0.65,4.96	0.24,4.30	0.00,3.73	0.00,3.43	0.00,2.96
6	3.85,8.83	3.35,8.34	2.85,7.83	2.35,7.33	1.86,6.82	1.38,6.27	0.93,5.69	0.51,5.06	0.12,1.38	0.00.3.64
7	4.65,10.03	4.15,9.53	3.65,9.03	3.15,8.53	2.66,8.02	2.16,7.51	1.69,6.97	1.23,6.40	0.80,5.79	0.01,4.45
8	5.47,11.21	4.97,10.71	4.47,10.21	3.97,9.71	3.47,9.21	2.98,8.70	2.48,8.19	2.00,7.66	1.54,7.10	0.68,5.88
9	6.30,12.37	5.80,11.88	5.30,11.38	4.80,10.88	4.30,10.38	3.80,9.87	3.30,9.37	2.81,8.86	2.33,8.33	1.47,7.21
10	7.14,13.54	6.64,13.04	6.14,12.54	5.64,12.04	5.14,11.54	4.64,11.04	4.14,10.53	3.64,10.03	3.15,9.52	2.19,8.46
11	7.99,14.69	7.49,14.19	6.99,13.69	6.49,13.19	5.99,12.69	5.49,12.19	4.99,11.69	4.49,11.19	3.99,10.69	3.01,9.66
12	8.84,15.83	8.34,15.33	7.84,14.83	7.34,14.33	6.84,13.83	6.34,13.33	5.84,12.84	5.34,12.33	4.84,11.83	3.85,10.82
13	9.70,16.98	9.20,16.47	8.70,15.97	8.20,15.47	7.70,14.97	7.20,14.47	6.70,13.97	6.20,13.47	5.70,12.97	4.70,11.97
14	10.56,18.11	10.06,17.61	9.56,17.11	9.07,16.61	8.57,16.11	8.06,15.61	7.57,15.11	7.07,14.61	6.56,14.11	5.57,13.11
15	11.43,19.24	10.94,18.74	10.43,18.24	9.93,17.74	9.44,17.54	8.94,16.74	8.43,16.24	7.93,15.74	7.43,15.24	6.43,14.24
16	12.31,20.37	11.81,19.86	11.31,19.36	10.81,18.87	10.31,18.36	9.81,17.87	9.31,17.37	8.81,16.86	8.31,16.37	7.31,15.37
17	13.18,21.49	12.68,20.99	12.18,20.49	11.69,19.99	11.19,19.49	10.68,18.98	10.18,18.49	9.69,17.99	9.19,17.49	8.19,16.49
18	14.07,22.61	13.57,22.10	13.07,21.60	12.57,21.11	12.07,20.61	11.57,20.11	11.07,19.60	10.57,19.11	10.07,18.61	9.07,17.61
19	14.95,23.72	14.45,23.22	13.95,22.72	13.45,22.22	12.95,21.72	12.45,21.22	11.95,20.72	11.45,20.22	10.95,19.72	9.95,18.72
20	15.84,24.83	15.34,24.33	14.84,23.83	14.34,23.33	13.84,22.83	13.34,22.34	12.84,21.83	12.34,21.33	11.84,20.83	10.84,19.84

表 23.1∼ 表 23.3 贝叶斯方法求得的信号区间内信号泊松事例数期望值 s 的最大后验密度信度区间 (先验分布为均匀分布). 其中 n 和 b 分别是信号区间内观测到的事例总数和已知平均本底事例数.

续表

$n\backslash b$	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0
0	0.11,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15	0.00,1.15
1	0.00,1.32	0.00,1.30	0.00,1.28	0.00,1.27	0.00,1.26	0.00.1.25	0.00,1.24	0.00,1.23	0.00,1.23	0.00,1.22
2	0.00,1.54	0.00,1.48	0.00,1.44	0.00,1.41	0.00,1.38	0.00,1.36	0.00,1.34	0.00,1.33	0.00,1.32	0.00,1.30
3	0.00,1.82	0.00,1.72	0.00,1.64	0.00,1.58	0.00,1.53	0.00,1.50	0.00,1.47	0.00,1.44	0.00,1.42	0.00,1.40
4	0.00,2.17	0.00,2.00	0.00,1.88	0.00,1.79	0.00,1.71	0.00,1.65	0.00,1.61	0.00,1.57	0.00,1.53	0.00,1.50
5	0.00,2.61	0.00,2.36	0.00,2.18	0.00,2.04	0.00,1.93	0.00,1.84	0.00,1.77	0.00,1.71	0.00,1.67	0.00,1.62
6	0.00,3.16	0.00,2.81	0.00.2.54	0.00,2.34	0.00,2.19	0.00,2.06	0.00,1.97	0.00,1.89	0.00,1.82	0.00,1.76
7	0.00,3.83	0.00,3.35	0.00,2.99	0.00,2.71	0.00,2.50	0.00,2.33	0.00,2.19	0.00,2.09	0.00,2.00	0.00,1.92
8	0.00,4.61	0.00,4.01	0.00,3.53	0.00,3.16	0.00,2.88	0.00,2.65	0.00,2.47	0.00,2.32	0.00,2.20	0.00,2.10
9	0.57,5.96	0.00,4.77	0.00,4.18	0.00,3.70	0.00,3.33	0.00,3.03	0.00,2.79	0.00,2.60	0.00,2.45	0.00,2.32
10	1.29,7.31	0.48,6.04	0.00,4.92	0.00,4.34	0.00,3.86	0.00,3.48	0.00,3.18	0.00,2.93	0.00,2.73	0.00,2.56
11	2.06,8.58	1.19,7.40	0.39,6.10	0.00,5.06	0.00,4.49	0.00,4.01	0.00,3.63	0.00,3.32	0.00,3.06	0.00,2.85
12	2.88,9.78	1.95,8.68	1.09,7.48	0.31,6.16	0.00,5.20	0.00,4.63	0.00,4.16	0.00,3.77	0.00,3.45	0.00,3.19
13	3.71,10.95	2.75,9.90	1.84,8.77	1.00,7.55	0.32,6.22	0.00,5.34	0.00,4.77	0.00,4.30	0.00,3.91	0.00,3.58
14	4.57,12.10	3.59,11.07	2.64,10.00	1.74,8.86	0.92,7.62	0.16,6.28	0.00,5.47	.00,4.91	0.00,4.44	0.00,4.04
15	5.44,13.24	4.44,12.22	3.47,11.19	2.53,10.10	1.65,8.94	0.84,7.68	0.10,6.34	0.00,5.59	0.00,5.04	0.00,4.57
16	6.31,14.36	5.31,13.36	4.33,12.34	3.36,11.29	2.43,10.19	1.56,9.01	0.76,7.74	0.03,6.39	0.00,5.72	0.00,5.17
17	7.18,15.49	6.18,14.48	5.20,13.48	4.21,12.45	3.26,11.39	2.34,10.27	1.48,9.08	0.70,7.80	0.00,6.47	0.00,5.84
18	8.07,16.61	7.07,15.61	6.07,14.60	5.08,13.59	4.10,12.56	3.16,11.48	2.25,10.35	1.41,9.14	0.63,7.85	0.00,6.58
19	8.95,17.72	7.95,16.72	6.95,15.72	5.96,14.71	4.97,13.69	4.00,12.65	3.06,11.57	2.17,10.43	1.34,9.20	0.57,7.91
20	9.84,18.84	8.84,17.83	7.84,16.83	6.84,15.83	5.85,14.82	4.86,13.80	3.90,12.75	2.97,11.66	2.09,10.50	1.27,9.26

表 23.2　信度水平 90%

$n\backslash b$	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0
0	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30
1	0.08,3.95	0.00,3.50	0.00,3.27	0.00,3.11	0.00,2.99	0.00,2.91	0.00,2.84	0.00,2.78	0.00,2.74	0.00,2.67
2	0.45,5.45	0.00,4.82	0.00,4.43	0.00,4.11	0.00,3.87	0.00,3.67	0.00,3.52	0.00,3.39	0.00,3.29	0.00,3.12
3	0.95,6.91	0.45,6.40	0.00,5.70	0.00,5.28	0.00,4.92	0.00,4.62	0.00,4.36	0.00,4.15	0.00,3.97	0.00,3.68
4	1.52,8.33	1.01,7.84	0.53,7.29	0.08,6.60	0.00,6.08	0.00,5.69	0.00,5.34	0.00,5.04	0.00,4.78	0.00,4.36
5	2.14,9.70	1.63,9.21	1.14,8.71	0.65,8.15	0.22,7.48	0.00,6.85	0.00,6.44	0.00,6.06	0.00,5.72	0.00,5.15
6	2.79,11.05	2.29,10.55	1.79,10.05	1.29,9.54	0.81,8.98	0.37,8.34	0.00,7.60	0.00,7.16	0.00,6.76	0.00,6.06
7	3.47,12.36	2.97,11.86	2.47,11.37	1.97,10.86	1.48,10.35	1.00,9.80	0.55,9.18	0.15,8.46	0.00,7.88	0.00,7.07
8	4.18,13.66	3.67,13.16	3.17,12.66	2.67,12.16	2.17,11.66	1.68,11.14	1.20,1.60	0.75,9.99	0.33,9.31	0.00,8.15
9	4.90,14.93	4.39,14.43	3.90,13.93	3.39,13.43	2.89,12.94	2.40,12.43	1.90,11.92	1.42,11.38	0.96,10.79	0.15,9.41
10	5.63,16.19	5.13,15.69	4.63,15.20	4.13,14.70	3.63,14.20	3.13,13.70	2.63,13.19	2.14,12.68	1.66,12.14	0.75,10.95
11	6.38,17.44	5.88,16.95	5.38,16.44	4.88,15.95	4.38,15.45	3.88,14.95	3.38,14.44	2.88,13.94	2.39,13.43	1.43,12.34
12	7.14,18.68	6.64,18.18	6.14,17.68	5.65,17.18	5.14,16.69	4.64,16.18	4.14,15.69	3.64,15.18	3.14,14.68	2.16,13.64
13	7.91,19.91	7.41,19.41	6.91,18.91	6.41,18.41	5.91,17.92	5.41,17.41	4.91,16.91	4.41,16.42	3.91,15.91	2.92,14.90
14	8.68,21.13	8.19,20.63	7.69,20.13	7.19,19.63	6.69,19.14	6.19,18.63	5.68,18.13	5.19,17.64	4.69,17.13	3.69,16.13
15	9.47,22.35	8.97,21.85	8.47,21.35	7.97,20.85	7.47,20.35	6.97,19.85	6.47,19.35	5.97,18.85	5.47,18.35	4.47,17.35
16	10.36,23.55	9.77,23.05	9.26,22.55	8.77,22.05	8.26,21.55	7.76,21.05	7.27,20.56	6.76,20.05	6.26,19.55	5.26,18.55
17	11.06,24.75	10.57,24.25	10.06,23.75	9.57,23.25	9.06,22.75	8.56,22.25	8.06,21.75	7.56,21.26	7.06,20.75	6.08,19.75
18	11.87,25.95	11.37,25.45	10.87,24.95	10.37,24.45	9.87,23.95	9.37,23.45	8.87,22.95	8.37,22.45	7.87,21.95	6.87,20.95
19	12.68,27.14	12.18,26.64	11.68,26.14	11.18,25.64	10.68,25.14	10.18,24.64	9.68,24.14	9.18,23.64	8.68,23.14	7.68,22.14
20	13.50,28.32	13.00,27.82	12.49,27.32	11.99,26.82	11.49,26.32	10.99,25.82	10.49,25.33	9.99,24.82	9.49,24.32	8.49,23.32

续表

$n\backslash b$	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0
0	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30	0.00,2.30
1	0.00,2.62	0.00,2.58	0.00,2.55	0.00,2.53	0.00,2.51	0.00,2.49	0.00,2.48	0.00,2.46	0.00,2.45	0.00,2.44
2	0.00,3.00	0.00,2.92	0.00,2.85	0.00,2.79	0.00,2.74	0.00,2.71	0.00,2.67	0.00,2.65	0.00,2.62	0.00,2.60
3	0.00,3.48	0.00,3.32	0.00,3.20	0.00,3.10	0.00,3.02	0.00,2.96	0.00,2.90	0.00,2.86	0.00,2.82	0.00,2.78
4	0.00,4.04	0.00,3.80	0.00,3.61	0.00,3.46	0.00,3.34	0.00,3.24	0.00,3.16	0.00,3.09	0.00,3.03	0.00,2.98
5	0.00,4.71	0.00,4.37	0.00,4.10	0.00,3.89	0.00,3.71	0.00,3.57	0.00,3.46	0.00,3.36	0.00,3.27	0.00,3.20
6	0.00,5.49	0.00,5.04	0.00,4.67	0.00,4.38	0.00,4.15	0.00,3.95	0.00,3.80	0.00,3.66	0.00,3.55	0.00,3.45
7	0.00,6.37	0.00,5.80	0.00,5.34	0.00,4.96	0.00,4.65	0.00,4.40	0.00,4.18	0.00,4.01	0.00,3.86	0.00,3.74
8	0.00,7.35	0.00,6.67	0.00,6.09	0.00,5.62	0.00,5.32	0.00,4.90	0.00,4.63	0.00,4.41	0.00,4.22	0.00,4.06
9	0.00,8.40	0.00,7.62	0.00,6.95	0.00,6.37	0.00,5.89	0.00,5.48	0.00,5.15	0.00,4.86	0.00,4.62	0.00,4.42
10	0.00,9.51	0.00,8.65	0.00,7.88	0.00,7.21	0.00,6.63	0.00,6.14	0.00,5.73	0.00,5.38	0.00,5.08	0.00,4.83
11	0.57,11.08	0.00,9.73	0.00,8.89	0.00,8.13	0.00,7.46	0.00,6.88	0.00,6.39	0.00,5.96	0.00,5.60	0.00,5.30
12	1.23,12.51	0.40,11.19	0.00,9.95	0.00,9.11	0.00,8.36	0.00,7.70	0.00,7.12	0.00,6.62	0.00,6.19	0.00,5.82
13	1.95,13.84	1.05,12.66	0.26,11.28	0.00,10.16	0.00,9.33	0.00,8.59	0.00,7.93	0.00,7.35	0.00,6.84	0.00,6.41
14	2.70,15.10	1.75,14.02	0.88,12.79	0.12,11.36	0.00,10.36	0.00,9.55	0.00,8.81	0.00,8.15	0.00,7.57	0.00,7.06
15	3.48,16.34	2.50,15.30	1.57,14.17	0.73,12.90	0.00,11.43	0.00,10.56	0.00,9.75	0.00,9.02	0.00,8.36	0.00,7.78
16	4.27,17.55	3.27,16.53	2.31,15.47	1.40,14.31	0.59,13.00	0.00,11.61	0.00,10.75	0.00,9.95	0.00,9.22	0.00,8.57
17	5.06,18.76	4.07,17.75	3.08,16.72	2.13,15.64	1.25,14.44	0.46,13.09	0.00,11.79	0.00,10.94	0.00,10.15	0.00,9.42
18	5.87,19.95	4.87,18.95	3.88,17.93	2.90,16.89	1.97,15.78	1.10,14.56	0.34,13.17	0.00,11.97	0.00,11.12	0.00,10.34
19	6.68,21.14	5.68,20.14	4.68,19.13	3.69,18.11	2.73,17.06	1.81,15.92	0.97,14.66	0.23,13.25	0.00,12.14	0.00,11.30
20	7.49,22.32	6.49,21.32	5.50,20.32	4.50,19.31	3.52,18.29	2.56,17.21	1.66,16.05	0.84,14.76	0.12,12.32	0.00,12.31

表 23.3　信度水平 95%

$n\backslash b$	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0
0	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00
1	0.04,4.75	0.00,4.35	0.00,4.11	0.00,3.94	0.00,3.81	0.00,3.72	0.00,3.64	0.00,3.58	0.00,3.53	0.00,3.45
2	0.31,6.33	0.00,5.77	0.00,5.38	0.00,5.06	0.00,4.80	0.00,4.60	0.00,4.43	0.00,4.29	0.00,4.17	0.00,3.99
3	0.73,7.87	0.23,7.36	0.00,6.75	0.00,6.34	0.00,5.97	0.00,5.66	0.00,5.39	0.00,5.16	0.00,4.97	0.00,4.66
4	1.22,9.38	0.72,8.89	0.24,8.31	0.00,7.67	0.00,7.23	0.00,6.83	0.00.6.48	0.00,6.16	0.00.5.89	0.00,5.43
5	1.77,10.83	1.27,10.34	0.77,9.83	0.30,9.22	0.00,8.53	0.00,8.08	0.00,7.66	0.00,7.27	0.00,6.92	0.00,6.33
6	2.36,12.23	1.86,11.74	1.36,11.25	0.86,10.72	0.40,10.12	0.00,9.36	0.00,8.90	0.00,8.46	0.00,8.05	0.00,7.33
7	2.98,13.61	2.48,13.12	1.98,12.62	1.48,12.12	0.99,11.59	0.53,10.99	0.11,10.27	0.00,9.70	0.00,9.25	0.00,8.42
8	3.63,14.96	3.13,14.47	2.63,13.97	2.13,13.47	1.63,12.96	1.14,12.44	0.67,11.85	0.25,11.16	0.00,10.48	0.00,9,57
9	4.30,16.29	3.80,15.79	3.30,15.30	2.79,14.80	2.29,14.30	1.80,13.79	1.31,13.26	0.84,12.69	0.41,12.03	0.00,10.77
10	4.98,17.60	4.48,17.10	3.98,16.61	3.48,16.11	2.98,15.61	2.48,15.11	1.98,14.60	1.50,14.07	1.02,13.51	0.18,12.15
11	5.68,18.90	5.18,18.40	4.68,17.90	4.18,17.40	3.68,16.91	3.16,16.40	2.68,15.90	2.19,15.40	1.70,14.87	0.77,13.71
12	6.40,20.18	5.90,19.69	5.40,19.19	4.90,18.69	4.40,18.19	3.90,17.69	3.40,17.19	2.90,16.69	2.40,16.18	1.43,15.11
13	7.12,21.46	6.62,20.96	6.12,20.46	5.62,19.96	5.12,19.46	4.62,18.96	4.12,18.46	3.62,17.96	3.12,17.46	2.14,16.44
14	7.86,22.72	7.36,22.22	6.86,21.72	6.36,21.22	5.86,20.72	5.36,20.22	4.86,19.72	4.36,19.22	3.86,18.72	2.86,17.72
15	8.60,23.97	8.11,23.47	7.61,22.98	7.10,22.47	6.61,21.98	6.11,21.48	5.60,20.98	5.10,20.48	4.60,19.98	3.60,18.98
16	9.36,25.22	8.86,24.75	8.36,24.22	7.86,23.72	7.36,23.22	6.86,22.72	6.36,22.22	5.86,21.72	5.36,21.22	4.36,20.22
17	10.12,26.46	9.62,25.96	9.12,25.46	8.62,24.96	8.12,24.46	7.62,23.96	7.12,23.46	6.62,22.96	6.12,22.46	5.12,21.46
18	10.89,27.69	10.39,27.19	9.89,26.69	9.39,26.19	8.89,25.69	8.39,25.19	7.89,24.69	7.39,24.19	6.88,23.69	5.89,22.69
19	11.66,28.92	11.16,28.41	10.66,27.91	10.16,27.41	9.66,26.92	9.16,26.42	8.66,25.92	8.16,25.42	7.66,24.92	6.66,23.92
20	12.44,30.14	11.94,29.64	11.44,29.14	10.94,28.64	10.44,28.14	9.94,27.64	9.44,27.14	8.94,26.63	8.44,26.14	7.44,25.14

续表

$n\backslash b$	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0
0	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00	0.00,3.00
1	0.00,3.39	0.00,3.34	0.00,3.31	0.00,3.28	0.00,3.25	0.00,3.23	0.00,3.22	0.00,3.20	0.00,3.19	0.00,3.18
2	0.00,3.86	0.00,3.75	0.00,3.67	0.00,3.60	0.00,3.55	0.00,3.50	0.00,3.46	0.00,3.43	0.00,3.40	0.00,3.38
3	0.00,4.42	0.00,4.24	0.00,4.10	0.00,3.99	0.00,3.89	0.00,3.81	0.00,3.75	0.00,3.69	0.00,3.64	0.00,3.60
4	0.00,5.09	0.00,4.82	0.00,4.60	0.00,4.43	0.00,4.29	0.00,4.17	0.00.4.07	0.00,3.98	0.00,3.91	0.00,3.85
5	0.00,5.85	0.00,5.48	0.00,5.18	0.00,4.94	0.00,4.73	0.00,4.57	0.00,4.43	0.00,4.31	0.00,4.22	0.00,4.13
6	0.00,6.73	0.00,6.24	0.00,5.84	0.00,5.52	0.00,5.25	0.00,5.03	0.00,4.85	0.00,4.69	0.00,4.55	0.00,4.44
7	0.00,7.70	0.00,7.10	0.00,6.60	0.00,6.18	0.00,5.84	0.00,5.55	0.00,5.31	0.00,5.11	0.00,4.94	0.00,4.78
8	0.00,8.76	0.00,8.05	0.00,7.45	0.00,6.94	0.00,6.51	0.00,6.15	0.00,5.84	0.00,5.58	0.00,5.36	0.00,5.17
9	0.00,9.89	0.00,9.09	0.00,8.38	0.00,7.77	0.00,7.25	0.00,6.81	0.00,6.44	0.00,6.12	0.00,5.84	0.00,5.61
10	0.00,11.05	0.00,10.18	0.00,9.39	0.00,8.69	0.00,8.08	0.00,7.56	0.00,7.11	0.00,6.72	0.00,6.38	0.00,6.10
11	0.00,12.25	0.00,11.33	0.00,10.47	0.00,9.69	0.00,8.99	0.00,8.38	0.00,7.85	0.00,7.39	0.00,6.99	0.00,6.64
12	0.55,13.87	0.00,12.50	0.00,11.59	0.00,10.74	0.00,9.97	0.00,9.27	0.00,8.66	0.00,8.12	0.00,7.65	0.00,7.25
13	1.19,15.32	0.35,13.99	0.00,12.75	0.00,11.85	0.00,11.01	0.00,10.24	0.00,9.55	0.00,8.93	0.00,8.39	0.00,7.91
14	1.89,16.67	0.97,15.50	0.18,14.09	0.00,12.99	0.00,12.10	0.00,11.26	0.00,10.50	0.00,9.81	0.00,9.20	0.00,8.65
15	2.61,17.96	1.65,16.88	0.77,15.65	0.02,14.17	0.00,13.22	0.00,12.34	0.00,11.51	0.00,10.75	0.00,10.07	0.00,9.45
16	3.36,19.22	2.38,18.18	1.44,17.07	0.59,15.78	0.00,14.37	0.00,13.45	0.00,12.57	0.00,11.75	0.00,11.00	0.00,10.31
17	4.12,20.46	3.13,19.44	2.16,18.39	1.24,17.23	0.43,15.89	0.00,14.59	0.00,13.67	0.00,12.80	0.00,11.99	0.00,11.24
18	4.89,21.69	3.89,20.69	2.90,19.66	1.94,18.59	1.05,17.38	0.27,15.98	0.00,14.79	0.00,13.88	0.00,13.02	0.00,12.21
19	5.66,22.92	4.66,21.92	3.66,20.91	2.69,19.87	1.75,18.76	0.88,17.51	0.14,16.07	0.00,15.00	0.00,14.10	0.00,13.24
20	6.44,24.14	5.44,23.14	4.44,23.13	3.45,21.12	2.48,20.06	1.56,18.92	0.72,17.63	0.01,16.14	0.00,15.20	0.00,14.30

示 例 索 引

例 1.1 继电器网路 1.4
例 1.2 契仑科夫计数器的探测效率 1.4
例 1.3 π° 介子的探测 1.4
例 1.4 粒子束流沾污和 δ 射线 1.4
例 1.5 扫描效率 (1) 1.4
例 1.6 pp 反应末态的分类 1.5
例 1.7 摸线币试验 1.5
例 2.1 $e^+e^- \to \mu^+\mu^-$ 反应的角分布 (1) 2.2
例 2.2 $Y = X^2$ 的概率分布 2.3
例 3.1 两个相互独立的 [0, 1] 区间均匀分布的卷积 3.4
例 3.2 达里兹图 3.7
例 3.3 放射性测量的误差 3.9
例 3.4 直角坐标测定值和极坐标测定值间的误差转换 3.9
例 4.1 探测器的探测效率 4.1
例 4.2 粒子反应产物的前后不对称性 (1) 4.1
例 4.3 试验次数的确定 4.1
例 4.4 扫描效率 (2) 4.1
例 4.5 粒子径迹探测效率 4.1
例 4.6 事例直方图 (1) 4.2
例 4.7 泡室中粒子径迹的气泡数目的分布 4.3
例 4.8 放射性衰变规律 4.3
例 4.9 放射源和本底辐射的叠加 4.3
例 4.10 粒子探测器计数的分布 4.4
例 4.11 计数时间间隔的细分 4.4
例 4.12 粒子反应产物的前后不对称性 (2) 4.4
例 4.13 事例直方图 (2) 4.4
例 4.14 云室中沿粒子径迹液滴数的分布 4.5
例 4.15 多丝室的空间分辨 4.7
例 4.16 泡室中粒子径迹气泡间的距离分布 4.8
例 4.17 两次相继的核衰变之间时间间隔的分布 4.8
例 4.18 成批数据的计算机在线处理 4.9
例 4.19 包含共振峰的不变质量谱 4.20.3

例 5.1　正态随机数的产生　5.2
例 7.1　总体均值已知时, 其方差的无偏估计量　7.3
例 7.2　总体均值未知时, 其方差的无偏估计量　7.3
例 7.3　三阶中心矩的无偏估计量　7.3
例 7.4　二项分布总体参数的有效无偏估计量　7.4
例 7.5　泊松总体均值的有效无偏估计量　7.4
例 7.6　正态总体均值的估计量　7.4
例 7.7　正态总体方差和标准差的估计量　7.4
例 7.8　柯西分布总体中参数不存在有效估计量和充分估计量　7.5
例 7.9　泊松总体中参数的充分、有效估计量　7.5
例 7.10　正态总体参数的充分统计量　7.5
例 7.11　正态总体 $N(\mu, \sigma^2)$ 的 μ, σ^2 联合充分统计量　7.5
例 7.12　指数分布参数的置信区间　7.6
例 7.13　束流动量的置信区间　7.7
例 8.1　不稳定粒子平均寿命的估计值　8.1
例 8.2　柯西分布中参数的估计　8.1
例 8.3　泊松分布中参数的估计　8.1
例 8.4　(0, 1) 分布中的参数估计　8.1
例 8.5　均匀分布中的参数估计　8.1
例 8.6　粒子平均寿命的极大似然估计量的渐近正态性质　8.3
例 8.7　粒子平均寿命极大似然估计量的方差　8.4.1
例 8.8　加权平均的方差　8.4.2
例 8.9　正态总体 $N(\mu, \sigma^2)$ 中 μ 和 σ^2 的极大似然估计量的方差　8.4.2
例 8.10　反质子极化实验 (1)　8.4.3
例 8.11　扫描效率 (3)　8.5.1
例 8.12　不稳定粒子平均寿命的置信区间　8.6.2
例 8.13　正态总体 $N(\mu, \sigma^2)$ 中参数 μ, σ^2 的联合似然域　8.6.3
例 8.14　同一物理量多次测量的合并估计　8.8
例 8.15　多个物理量不对称误差的合并估计　8.8
例 8.16　实验分辨对极大似然估计的影响　8.9
例 9.1　直线拟合 (1)　9.2
例 9.2　抛物线拟合　9.2.3
例 9.3　两个独立的实验结果的合并　9.2.3
例 9.4　直线拟合 (2)　9.2.4
例 9.5　带电粒子的螺旋线径迹参数　9.3
例 9.6　圆弧径迹的拟合　9.3
例 9.7　直线拟合 (3)　9.4.1
例 9.8　反质子极化实验 (2)　9.5

例 9.9 粒子角动量分析 (1) 9.5
例 9.10 通过固定点的直线拟合 9.7
例 9.11 三角形的角 9.7
例 9.12 V^0 事例的运动学分析 (1) 9.8.1
例 9.13 强作用常数 α_S 的合并估计 9.10
例 10.1 反质子极化实验 (3) 10.3
例 10.2 粒子角动量分析 (2) 10.3
例 10.3 矩法应用于多个实验结果的合并 10.3
例 12.1 单个 π^0 和多个 π^0 事例的区分 12.1.2
例 12.2 粒子平均寿命的奈曼 —— 皮尔逊检验 12.2.1
例 12.3 正态分布均值的似然比检验 12.2.2
例 12.4 粒子平均寿命的似然比检验 12.2.2
例 12.5 粒子衰变理论的似然比检验 12.2.2
例 12.6 两个测量装置测量结果的比较 12.3.3
例 12.7 Ω^- 超子质量 12.3.4
例 12.8 $e^+e^- \to \mu^+\mu^-$ 反应的角分布 (2) 12.4.4
例 12.9 $e^+e^- \to \mu^+\mu^-$ 反应总截面随能量的分布 12.4.4
例 12.10 V^0 事例的运动学分析 (2) 12.4.4
例 12.11 低统计实验的拟合优度检验 12.4.5
例 12.12 动量分量的独立性 12.6
例 12.13 体重与肺活量的相关性 12.7.3
例 12.14 鸟的翼长、体长和嘴长间的相关性 12.7.4
例 12.15 汽车发动机耗油量 (1) 12.8.1
例 12.16 两组不变质量测量值之间的一致性 12.8.2
例 12.17 实验直方图与理论分布的比较 12.8.3
例 12.18 π^0 寿命两组测量值的一致性检验 12.8.5
例 12.19 汽车发动机耗油量 (2) 12.8.5
例 12.20 多组实验数据之间的一致性 12.8.7
例 13.1 用贝叶斯假设求正态分布均值的后验分布 13.2.3
例 13.2 正态分布参数的不变先验分布 13.2.3
例 13.3 指数分布参数的不变先验分布 13.2.3
例 13.4 指数分布参数的 Jeffreys 先验分布 13.2.3
例 13.5 二项分布参数的 Jeffreys 先验分布 13.2.3
例 13.6 泊松分布参数的 Jeffreys 先验分布 13.2.3
例 13.7 正态分布参数的 Jeffreys 先验分布 13.2.3
例 13.8 二项分布参数的共轭先验分布 13.2.3
例 13.9 几何分布参数的共轭先验分布 13.2.3
例 13.10 泊松分布参数的共轭先验分布 13.2.3

例 13.11　指数分布参数的共轭先验分布　13.2.3
例 13.12　正态分布参数的共轭先验分布　13.2.3
例 13.13　探测效率的贝叶斯后验期望估计与极大似然估计　13.3.2
例 13.14　探测效率的经验贝叶斯估计　13.3.3
例 13.15　正态总体均值的信度区间　13.3.4
例 13.16　本底分布上叠加小信号情况下信号事例数的确定　13.3.4
例 13.17　正态总体均值的贝叶斯假设检验　13.3.5
例 14.1　蒙特卡罗方法计算积分 $I=\int_0^1 \mathrm{e}^x \mathrm{d}x$　14.4.4
例 15.1　共振产生的确定　15.7.2

《现代物理基础丛书》已出版书目

(按出版时间排序)

1. 现代声学理论基础	马大猷 著	2004.03
2. 物理学家用微分几何 (第二版)	侯伯元, 侯伯宇 著	2004.08
3. 数学物理方程及其近似方法	程建春 编著	2004.08
4. 计算物理学	马文淦 编著	2005.05
5. 相互作用的规范理论 (第二版)	戴元本 著	2005.07
6. 理论力学	张建树, 等 编著	2005.08
7. 微分几何入门与广义相对论 (上册 · 第二版)	梁灿彬, 周彬 著	2006.01
8. 物理学中的群论 (第二版)	马中骐 著	2006.02
9. 辐射和光场的量子统计	曹昌祺 著	2006.03
10. 实验物理中的概率和统计 (第二版)	朱永生 著	2006.04
11. 声学理论与工程应用	朱海潮, 等 编著	2006.05
12. 高等原子分子物理学 (第二版)	徐克尊 著	2006.08
13. 大气声学 (第二版)	杨训仁, 陈宇 著	2007.06
14. 输运理论 (第二版)	黄祖洽 著	2008.01
15. 量子统计力学 (第二版)	张先蔚 编著	2008.02
16. 凝聚态物理的格林函数理论	王怀玉 著	2008.05
17. 激光光散射谱学	张明生 著	2008.05
18. 量子非阿贝尔规范场论	曹昌祺 著	2008.07
19. 狭义相对论 (第二版)	刘 辽, 等 编著	2008.07
20. 经典黑洞与量子黑洞	王永久 著	2008.08
21. 路径积分与量子物理导引	侯伯元, 等 著	2008.09
22. 量子光学导论	谭维翰 著	2009.01
23. 全息干涉计量——原理和方法	熊秉衡, 李俊昌 编著	2009.01
24. 实验数据多元统计分析	朱永生 编著	2009.02
25. 微分几何入门与广义相对论(中册 · 第二版)	梁灿彬, 周彬 著	2009.03
26. 中子引发轻核反应的统计理论	张竞上 著	2009.03
27. 工程电磁理论	张善杰 著	2009.08
28. 微分几何入门与广义相对论(下册 · 第二版)	梁灿彬, 周彬 著	2009.08
29. 经典电动力学	曹昌祺 著	2009.08
30. 经典宇宙和量子宇宙	王永久 著	2010.04

31. 高等结构动力学(第二版)	李东旭 著	2010.09
32. 粉末衍射法测定晶体结构(第二版 · 上、下册)	梁敬魁 编著	2011.03
33. 量子计算与量子信息原理	Giuliano Benenti 等著	
——第一卷：基本概念	王文阁, 李保文 译	2011.03
34. 近代晶体学(第二版)	张克从 著	2011.05
35. 引力理论(上、下册)	王永久 著	2011.06
36. 低温等离子体	B. M. 弗尔曼, И. M. 扎什京 编著	
——等离子体的产生、工艺、问题及前景	邱励俭 译	2011.06
37. 量子物理新进展	梁九卿，韦联福 著	2011.08
38. 电磁波理论	葛德彪，魏 兵 著	2011.08
39. 激光光谱学	W. 戴姆特瑞德 著	
——第 1 卷：基础理论	姬 扬 译	2012.02
40. 激光光谱学	W. 戴姆特瑞德 著	
——第 2 卷：实验技术	姬 扬 译	2012.03
41. 量子光学导论(第二版)	谭维翰 著	2012.05
42. 中子衍射技术及其应用	姜传海, 杨传铮 编著	2012.06
43. 凝聚态、电磁学和引力中的多值场论	H. 克莱纳特 著	
	姜 颖 译	2012.06
44. 反常统计动力学导论	包景东 著	2012.06
45. 实验数据分析(上册)	朱永生 著	2012.06
46. 实验数据分析(下册)	**朱永生 著**	**2012.06**